

ON GOVERNANCE

COMMENTARY

What Travels With a Sequence? Provenance and Governance in the Next Phase of Digital Sequence Information

By **June Rubis** and **Anisha Buda**

The digital sequence information (DSI) Multilateral Mechanism allows monetary benefit-sharing without tracing every sequence to every downstream use. But that does not make provenance, knowledge relationships, or Indigenous and community governance irrelevant. The next phase of DSI implementation raises a different question: What should remain connected to biological information as it moves through global data infrastructures?



Yarsagumba harvesting: A seasonal livelihood in the High Himalayas. Photograph by Aneeta Buda.

After SBI-7 in Nairobi, Anisha kept returning to yarsagumba

Yarsagumba (*Ophiocordyceps sinensis*) is found across the Himalayan and Tibetan regions, including Nepal, Bhutan, India, and China. In parts of Nepal, its seasonal harvest is deeply connected to household livelihoods and local economies. Research from the Dolpa district of Nepal has shown that income from yarsagumba significantly contributes to mountain households, while work across the Himalayan region has documented harvest intensity on wild populations and impacts of climate change (Hopping et al., 2018; Shrestha & Bawa, 2014). For Anisha, the connection between yarsagumba and livelihoods

is grounded in family and community experience. Her family has traded yarsagumba over many years. She remembers families whose livelihoods depended heavily on the harvest. She has experienced how these conditions have become harder as harvests decline and markets change.

These changes have multiple causes and should not be attributed to DSI. They nevertheless shaped the question Anisha carried away from Nairobi, where she participated in the seventh meeting of the Subsidiary Body on Implementation

DOI: <https://doi.org/10.70841/AH78DI>. © 2026 The Author. Published by the ICCA Consortium.

Published under the terms of the [Creative Commons Attribution-NonCommercial 4.0 International](https://creativecommons.org/licenses/by-nc/4.0/) license.

(SBI-7) of the Convention on Biological Diversity (CBD).

What happens when a species deeply embedded in particular livelihoods, histories, and knowledge systems also becomes available to science through information that can circulate independently of those relationships?



Yarsagumba harvested, cleaned, and ready to sell. Photograph by Aneeta Buda.

The yarsagumba example does not provide a straightforward claim concerning unauthorized access to genetic resources from Nepal. Research undertaken on material collected within another range state may never involve access to Nepalese genetic resources. Its significance for DSI lies elsewhere. Biological origin, scientific use, and social or economic consequence can occupy different geographies.

For Anisha, the question is also grounded in the way yarsagumba is lived and governed in Dolpa. Belonging to the Kaike Magar Indigenous Community, she grew up hearing about yarsagumba—the fascination of an insect that “turns into a plant,” as well as the way its harvest sustains families. In upper Dolpa, where winters are harsh and people may move to lower areas for the season, income from yarsagumba is crucial for getting households through the winter. It appears in the month of Baisakh (mid-April) and is harvested for around a month in Jestha.

Harvest is governed through relationships between communities and forests. Anisha describes that villages without their own harvesting forest pay royalties to Indigenous Peoples whose forests they enter, alongside fees associated with local government and protected areas. Harvesters also make collective decisions about when to pause collection so that yarsagumba has time to grow before harvesting resumes. For her, this is a governance system grounded in customary practice.

She also recalls concern about declining market prices and growing uncertainty among community around reports that China now has its own yarsagumba: “We don’t know how they are doing it, but we heard they have their own yarsagumba, which is why market price is so low.” Another observation she grew up hearing was that, despite generations of harvesting, yarsagumba continues to return. It indicates the significance of customary practices that have been going on for generations. These accounts carry ecological, economic, and governance knowledge that would never be visible in digital sequence information or physical genetic resources.

Why Yarsagumba Is a Useful DSI Example

Yarsagumba (*Ophiocordyceps sinensis*) is found across several Himalayan and Tibetan range states. Indigenous communities in Nepal have longstanding livelihood, knowledge, and trading relationships with the species, while sequencing or research may involve biological material collected elsewhere.

The example shows how biological origin, scientific use, and social or economic consequences can occupy different geographies. A genetic sequence from a specimen collected in one country could be used by researchers or companies in another; at the same time, communities elsewhere maintain longstanding livelihood and knowledge relationships with the same species. These overlapping relationships complicate the question of who should receive benefits from the use of that sequence and on what basis.

This example helps explain the rationale for multilateral benefit-sharing, while also raising questions about what happens to provenance, knowledge relationships, and governance as sequence information circulates independently.

The multilateral mechanism for benefit-sharing from DSI responds in part to this problem of circulation. It allows monetary benefits to be shared multilaterally without requiring every provider country to establish a direct relationship between a particular sequence and a particular commercial use (Convention on Biological Diversity [CBD], 2024; Scholz et al., 2022). The Cali

Fund gives this monetary component an institutional form and provides that, where appropriate and subject to national circumstances and legislation, at least half of its funding should support the self-identified needs of Indigenous Peoples and local communities, including women and youth, through government authorities or by direct payments through institutions identified

by Indigenous Peoples and local communities (CBD, 2024).

As implementation of DSI advances, the wider Multilateral Mechanism is also coming into view. SBI-7 brought databases, provenance, permissions, traditional knowledge, and data governance more directly into the negotiations. The draft recommendation going forward to the 17th meeting of the Conference of the Parties (COP17) of CBD contains proposals concerning work led by Indigenous Peoples and local

communities with database operators and scientific institutions, alongside possible measures on provenance, lawful submission, permissions, the CARE, FAIR, and TRUST principles, and the UNESCO Recommendation on Open Science (CBD, 2024, 2026).

These discussions raise a further question: **What happens to existing relationships of knowledge, Indigenous authority, and local community governance when biological information begins to circulate as DSI?**

From Access to Circulation

Access and benefit-sharing has developed law, policy, and Indigenous advocacy around the conditions under which genetic resources and associated traditional knowledge are accessed. Free, prior, and informed consent (FPIC), prior informed consent, mutually agreed terms, customary laws, and community protocols all address questions of authority and governance at important points in these relationships.

DSI creates a longer chain. A sequence deposited today may be reused years later by researchers or companies with no relationship to the original collector, provider country, or community. It may enter other databases, be combined with information from multiple origins, or acquire value through technologies that did not exist when the material was first collected. Work on future-proofing the Multilateral Mechanism has drawn particular attention to artificial intelligence and other technologies that can increase the distance between particular sequences, their origins, and later forms of use (Faggionato et al., 2024).

This temporal dimension raises questions about how authority expressed through FPIC and other forms of governance continues to matter as information moves through systems designed for circulation and reuse.

The monetary benefit-sharing governed through the Multilateral Mechanism can operate without reconstructing a bilateral payment relationship for every sequence and every

downstream user (Scholz et al., 2022). That multilateral design does not by itself displace separate rights or obligations that may arise under national Access and Benefit-sharing (ABS) measures, including in relation to Indigenous Peoples and local communities and associated traditional knowledge (CBD, 2024). Provenance and governance information have other functions. They can support scientific integrity and attribution, national ABS implementation, and recognition of traditional knowledge and the rights and interests of Indigenous Peoples and local communities (CBD, 2024; Raposo et al., 2023).

Some ABS relationships also depend on continuing forms of connection that are not reducible to monetary allocation, including research relationships, sharing of research outcomes, and access to resulting products and services (Ramakrishnan, 2026).

Tracing every downstream use through large datasets can be difficult and challenging, particularly when DSI is combined across multiple origins and undergoes computational transformation (Faggionato et al., 2024; Raposo et al., 2023). But the extent of that difficulty is also shaped by choices about database design, provenance requirements, and user accountability (Ramakrishnan, 2026). The question is how available provenance is being treated, what other governance information can remain connected to DSI, and what the multilateral system provides when those relationships become increasingly difficult to follow.

What Becomes Visible

Digital infrastructures are well suited to making some forms of information portable. A sequence can retain a scientific identifier, taxonomic information, and, where recorded, information concerning geographic origin.

Recent changes in database practice also show that more provenance information is now being recorded. The International Nucleotide Sequence Database Collaboration (INSDC) began introducing mandatory country/location and collection-date metadata for new submissions in 2023, extending the standard across the remaining sequence submission routes by the end of 2024, subject to specified exemptions (INSDC, 2023). A 2026 analysis found geographical collection information for 46% of sequences submitted to INSDC over the previous five years, compared with about 17% of the full corpus (Raposo et al., 2026). This suggests meaningful improvement in recent records alongside a substantial legacy gap.

However, the existence of a geographic field should not be equated with complete provenance. INSDC's /geo_loc_name qualifier records the locality where the sequenced sample was isolated or sampled (INSDC, 2024). That information can be important for science and relevant to ABS implementation, but it does not by itself establish the provider country in ABS terms, identify a source collection, an Indigenous territory, or community relationship, or record the authority and conditions under which material or knowledge was shared. Recent allocation modeling illustrates why this distinction matters: using INSDC sequence counts as a proxy for geographic origin can shift the measure towards sequencing and data infrastructure rather than the geographic origin of the underlying biological material (Makowska-Curran, 2026). The political meaning of apparently technical categories therefore depends in part on how they are operationalized. Current discussions could strengthen provenance and create ways of signaling permissions, traditional knowledge associations, and wider

data-governance considerations (CBD, 2026; Raposo et al., 2024).

Less readily carried are the relationships surrounding biological resources. Traditional knowledge and authority may be differentiated by gender, generation, lineage, livelihood, ritual responsibility, family, or territorial relationship. Some knowledge may be widely held; other knowledge may be restricted or governed by particular people or institutions. Describing information as “traditional knowledge associated with Indigenous Peoples and local communities” may therefore leave unresolved who holds that knowledge, who can make decisions concerning its disclosure or use, and what protocols apply. The Dolpa arrangements described earlier make the distinction concrete. A sequence record might retain a collection locality, but it would not tell a later user whose forest was entered, whose authority applied, or how harvesters decided collectively when harvesting should pause.

This points to a distinction between provenance and governance.

Provenance concerns the source and history through which biological material or information has traveled: where material was collected, from what source, and through which institutional or other relationships it entered scientific infrastructures. Governance concerns how decisions about

knowledge and its use are made, including who holds relevant authority and what protocols apply. More detailed metadata cannot by itself answer those governance questions.

The CARE Principles are useful because they bring Collective Benefit, Authority to Control, Responsibility, and Ethics into data governance (Carroll et al., 2020). Authority to Control is particularly relevant because governance is not exhausted by the initial decision to share data; it also concerns conditions of access, use, reuse, and stewardship across the data lifecycle. Work on Indigenous data sovereignty in genomics raises similar questions about whether open infrastructures adequately accommodate Indigenous rights and interests (Mc Cartney et al., 2022).

At the same time, better governance should not require increasingly detailed documentation of Indigenous knowledge to demonstrate that rights or responsibilities exist. In some circumstances, what needs to travel is not the knowledge itself, but an indication that a community protocol or governance relationship applies and that further engagement is required. Local Contexts provides one example: its Traditional Knowledge and Biocultural Labels allow communities to communicate provenance, protocols, and permissions without necessarily placing underlying knowledge into open circulation (Liggins et al., 2021; Local Contexts, n.d.).



Cattle grazing across the Himalayan highlands, sustaining traditional pastoral livelihoods. Photograph by Angad Kumar Buda.

What Database Governance Can Carry

The emerging SBI-7 text provides a practical entry point through proposals for Indigenous Peoples and local communities, including women and youth, to develop self-led materials for database operators, scientific institutions, and other actors (CBD, 2026). This should not be understood only as a need for communities to learn how DSI works. It also creates an opportunity for scientific and database infrastructures to encounter governance arrangements that may otherwise remain invisible: customary law, community protocols, kinship, ritual responsibility, territorial authority, and other forms through which decisions about knowledge and biological resources are made. Work from the DSI Scientific Network has examined geographical-origin information, database governance, and the technical challenges of maintaining relationships between DSI and its origins (Faggionato et al., 2024; Raposo et al., 2023, 2024). Nithin Ramakrishnan's digital-justice analysis draws attention to another part of the problem: how responsibility for legality and benefit-sharing may fall heavily on data providers and national authorities, while database operators assume more limited responsibility for uses enabled through their infrastructures (Ramakrishnan, 2026).

These perspectives point to a practical question of design. Database systems have technical constraints, but they also embody choices about where information and responsibility sit: which provenance fields are required, how missing information is handled, whether links remain connected as data move between infrastructures, and what responsibilities database operators and users carry when information is reused (Ramakrishnan, 2026; Raposo et al., 2024). The recently published IEEE 2890-2025 Recommended Practice for Provenance of Indigenous Peoples' Data offers a concrete point of reference. It establishes parameters for recording provenance

in ways intended to support Indigenous governance, decision-making, and engagement in present and future uses of Indigenous Peoples' data (IEEE, 2025). It does not resolve the DSI question, but it demonstrates that provenance can be designed with governance relationships in view.

Openness of scientific access and user accountability also need not be organized identically across all data infrastructures.

The practical questions ahead are therefore not how much metadata can be attached to a sequence. They include how provenance and community-defined governance information remain connected as data move between infrastructures, how databases signal the existence of traditional knowledge or relevant protocols, what happens when information from multiple origins is combined, and which forms of information should remain outside databases.

That final question also has a temporal dimension. Information that appears limited or non-sensitive when recorded may acquire new meanings or become more revealing when combined with other datasets or analytical technologies later (Carroll et al., 2020).

FAIR, CARE, and TRUST, together with the UNESCO Recommendation on Open Science, offer different starting points for these discussions (Carroll et al., 2020; CBD, 2026; UNESCO, 2021). But a universal Indigenous metadata template is unlikely to answer the governance question. The objective is not to make every governance relationship necessarily machine-readable. It may instead be to enable digital infrastructures to recognize that such relationships exist and create routes through which users can encounter them.

What Is at Stake at CBD COP17

The Cali Fund remains central to the DSI agenda for Indigenous Peoples and local communities. Its credibility will depend on whether resources are mobilized and on how the provision that, where appropriate and subject to national circumstances and legislation, at least half of the Fund should support the self-identified needs of Indigenous Peoples and local communities is operationalized: through what access pathways, how priorities are set, and how meaningful authority over delivery is preserved. These questions are distinct from the country-allocation methodology for funding to Parties (CBD, 2024).

But the rights question does not end with monetary allocation. The Multilateral Mechanism itself is without prejudice to national ABS measures and must respect the rights of Indigenous Peoples and local communities (CBD, 2024). Customary laws, community protocols, and other governance arrangements through which biological resources and

associated knowledge may have become accessible therefore remain relevant. A credible provider-side bargain can help preserve the political space in which biodiversity, provenance, and benefit-sharing remain consequential, particularly for developing countries advancing provider-side interests. That does not make provider-country interests and the interests of Indigenous Peoples and local communities identical: their distinct rights, authorities, and relationships require their own assurance.

The wider Multilateral Mechanism carries another part of the bargain.

As information moves further from the conditions in which biological resources and knowledge first became accessible, the system will need to determine what forms of provenance, responsibility, and governance remain visible.



Highland landscape of the yarsagumba harvesting area. Photograph by Aneeta Buda.

ANISHA I After SBI-7

“

I returned from SBI-7 thinking that Indigenous organizations in Nepal need to follow both the Cali Fund and the wider DSI debate. For me, the starting point is not specialist scientific expertise. Indigenous Peoples already have their own knowledge, responsibilities, and relationships with species and the lands where they live. DSI may be technically complex, but understanding what is at stake does not require a background in those fields: “You do not need to come from a science background to understand DSI.” The commitment that at least 50% of Cali Fund resources will support Indigenous Peoples and local communities makes the issue practical, but I understand that the implications of DSI can be much wider than the Fund. It concerns traditional knowledge, culture, and the relationships Indigenous Peoples have with species, as well as where sequence information travels and how it is used.

I am particularly concerned that DSI should not become a route through which information derived from genetic resources or associated Indigenous knowledge is commercially used without the relevant Indigenous Peoples understanding how that information is being used or agreeing to the conditions of use. Indigenous authority, in my view, is not just about being included in a process. It requires meaningful decision-making, FPIC where Indigenous rights and authority are engaged, and the ability to withhold consent. The CARE Principles—Collective Benefit, Authority to Control, Responsibility, and Ethics—are therefore an important guide for DSI governance.

”

A sequence can remain usable long after the original research relationship has ended. It can cross borders and institutions, enter new datasets, and acquire forms of value that could not have been anticipated at the time of collection. The national, territorial, and community relationships surrounding the biological material do not disappear when that happens. Nor does the authority through which knowledge or biological resources were shared.

As DSI moves through the global data infrastructures, COP17 has an opportunity to ask what should continue to travel with it and to what end, and whose authority must still be recognized when it does.

References

Carroll, S. R., Garba, I., Figueroa-Rodríguez, O. L., Holbrook, J., Lovett, R., Materechera, S., Parsons, M., Raseroka, K., Rodríguez-Lonebear, D., Rowe, R., Sara, R., Walker, J. D., Anderson, J., & Hudson, M. (2020). The CARE principles for Indigenous data governance. *Data Science Journal*, 19, Article 43. <https://doi.org/10.5334/dsj-2020-043>

Convention on Biological Diversity. (2024). Decision 16/2: *Digital sequence information on genetic resources (CBD/COP/DEC/16/2)*. <https://www.cbd.int/doc/decisions/cop-16/cop-16-dec-02-en.pdf>

- Convention on Biological Diversity. (2026). *Digital sequence information on genetic resources (CBD/SBI/7/L.13)*. <https://www.cbd.int/meetings/SBI-07>
- Faggionato, D., Orozco, P., & Scholz, A. H. (2024). *Future-proofing the DSI multilateral mechanism: Possible implications of artificial intelligence & other upcoming technologies*. DSI Scientific Network. <https://doi.org/10.5281/zenodo.14196246>
- Hopping, K. A., Chignell, S. M., & Lambin, E. F. (2018). The demise of caterpillar fungus in the Himalayan region due to climate change and overharvesting. *Proceedings of the National Academy of Sciences of the United States of America*, 115(45), 11489–11494. <https://doi.org/10.1073/pnas.1811591115>
- IEEE. (2025). IEEE 2890-2025: *IEEE recommended practice for provenance of Indigenous Peoples' data*. <https://standards.ieee.org/ieee/2890/10318/>
- International Nucleotide Sequence Database Collaboration. (2023). *INSDC spatiotemporal metadata – minimum standards update (03-03-2023)*. <https://www.insdc.org/news/insdc-spatiotemporal-metadata-minimum-standards-update-03-03-2023/>
- International Nucleotide Sequence Database Collaboration. (2024, July 8). *geo_loc_name* qualifier vocabulary. https://www.insdc.org/submitting-standards/geo_loc_name-qualifier-vocabulary/
- Liggins, L., Hudson, M., & Anderson, J. (2021). Creating space for Indigenous perspectives on access and benefit-sharing: Encouraging researcher use of the Local Contexts Notices. *Molecular Ecology*, 30(11), 2477–2482. <https://doi.org/10.1111/mec.15918>
- Local Contexts. (n.d.). *Labels*. Retrieved September 21, 2026, from <https://localcontexts.org/labels/about-the-labels/>
- Makowska-Curran, A. (2026). *From global biodiversity commitments to local action: Revenue potential and allocation dynamics of the Cali Fund* (Working Paper No. 450). Grantham Research Institute on Climate Change and the Environment, London School of Economics and Political Science. <https://doi.org/10.21953/researchonline.lse.ac.uk.00138657>
- Mc Cartney, A. M., Anderson, J., Liggins, L., Hudson, M. L., Anderson, M. Z., TeAika, B., Geary, J., Cook-Deegan, R., Patel, H. R., & Phillippy, A. M. (2022). Balancing openness with Indigenous data sovereignty: An opportunity to leave no one behind in the journey to sequence all of life. *Proceedings of the National Academy of Sciences of the United States of America*, 119(4), Article e2115860119. <https://doi.org/10.1073/pnas.2115860119>
- Ramakrishnan, N. (2026). *Access to digital sequence information (DSI) and fair and equitable sharing of benefits arising from its utilization: A digital justice perspective*. IT for Change; Third World Network. https://itforchange.net/sites/default/files/2645/Access_to_Digital_Sequence_Information_DSI.pdf
- Raposo, D. S., Cepeda, M. L., Ebert, B., & Scholz, A. H. (2023). *Challenges and opportunities of geographical origin information in DSI benefit-sharing: A global analysis from the academic sector*. DSI Scientific Network. <https://doi.org/10.5281/zenodo.10050454>
- Raposo, D. S., Orozco, P., McCartney, A., Freitag, J., Rouard, M., Ebert, B., & Scholz, A. H. (2024). *Digital sequence information (DSI) data governance practices to support benefit-sharing and science*. DSI Scientific Network. <https://doi.org/10.5281/zenodo.12755428>
- Raposo, D. S., Faggionato, D., Ebert, B., Orozco, P., Cochrane, G., Bao, Y., Arita, M., Muñoz-García, M., Freitag, J., Hirsch, T., Cook, C. E., Vos, R., Reimer, L., Ratnasingham, S., Rouard, M., Brinkley, S., Lange, M., Scholz, U., Weise, S., ... Scholz, A. H. (2026). How can biological databases support the new UN mechanism for benefit-sharing from digital sequence information? *Scientific Data*, 13, 971. <https://doi.org/10.1038/s41597-026-07725-y>
- Scholz, A. H., Freitag, J., Lyal, C. H. C., Sara, R., Cepeda, M. L., Cancio, I., Sett, S., Hufton, A. L., Abebaw, Y., Bansal, K., Benbouza, H., Boga, H. I., Brisse, S., Bruford, M. W., Clissold, H., Cochrane, G., Coddington, J. A., Deletoille, A.-C., García-Cardona, F., ... Overmann, J. (2022). Multilateral benefit-sharing from digital sequence information will support both science and biodiversity conservation. *Nature Communications*, 13, 1086. <https://doi.org/10.1038/s41467-022-28594-0>
- Shrestha, U. B., & Bawa, K. S. (2014). Economic contribution of Chinese caterpillar fungus to the livelihoods of mountain communities in Nepal. *Biological Conservation*, 177, 194–202. <https://doi.org/10.1016/j.biocon.2014.06.019>
- UNESCO. (2021). *Recommendation on open science*. <https://www.unesco.org/en/legal-affairs/recommendation-open-science>

About the authors

June Rubis is a Bidayuh scholar from Sarawak, Malaysian Borneo, a Macquarie University Research Fellow, and co-founder of Building Initiatives in Indigenous Heritage (BiIH). She works on questions of benefit-sharing, governance, and implementation across the wider Multilateral Mechanism, including direct access and distribution pathways for the Cali Fund. Anisha Buda is from the Kaike Magar Indigenous Community of Dolpa, Nepal. She is a Program Officer at the National Indigenous Women Forum (NIWF), a Secretariat Committee member of ICCA Network Nepal, and an Interim Committee Member of the International Indigenous-Led Fire Adaptation Network (IIFAN). Both supported the International Indigenous Forum on Biodiversity (IIFB) negotiations on digital sequence information at SBI-7 in Nairobi in 2026. This commentary is informed by that shared experience but reflects the authors' own analysis.

Reviewed by: Katielee Riddle, Jasmine Kindness and other peer reviewers who would wish to remain anonymous.

Copyediting: Mahatub Khan Badhon

Editorial Lead: Mohammad Arju

The author is deeply grateful to the peer reviewers whose thoughtful comments, challenges, and suggestions greatly strengthened this paper.

Correspondence concerning this article should be addressed to June Rubis at jrubis@gmail.com

The views expressed in this commentary are those of the author and do not necessarily reflect the views or policy positions of the ICCA Consortium.

Recommended citation: Rubis, J., & Buda, A. (2026). *What travels with a sequence? Provenance and governance in the next phase of digital sequence information* (On Governance). ICCA Consortium. <https://doi.org/10.70841/AH78DI>