



Downloaded from <https://www.science.org> on October 24, 2024

OPEN WITH CARE

Indigenous researchers and communities are reshaping how Western science thinks about open access to data

By **Sandeep Ravindran**

When Leslie “Leke” Hutchins presented his data at a scientific meeting in August 2023, he was met with some consternation. A University of California (UC), Berkeley Ph.D. student in conservation biology and a Kānaka ‘Ōiwi (Native Hawaiian), Hutchins had studied the diversity of invertebrates such as insects and spiders on Native Hawaiian coffee plantations. But after extensive discussions with the farmers, he redacted the species names and where they were collected,

information the farmers considered culturally sensitive. After Hutchins announced at the meeting that he would not be providing those data publicly, one attendee asked, “What do you mean you’re having some of your data closed off? You know, that’s not fair,” Hutchins recalls.

In academia, fairness has increasingly come to mean research data are publicly

shared and widely accessible. For example, the Ecological Society of America (ESA), at whose annual meeting Hutchins was presenting, follows an open research policy for its journals, which requires all underlying data in accepted manuscripts to be “made available in a permanent, publicly accessible data archive or repository ... with rare exceptions.”

But Hutchins had good reason to protect his data, based on a different principle of fairness: the right of Indigenous people to control how data about them, their lands, and their cultures are collected and used. He’s part of a growing movement among Indig-

GLOBAL EQUITY IN SCIENCE

S This story is part of an ongoing News series. Read more at scim.ag/GlobalEquity

PHOTO: LEKE HUTCHINS



Leke Hutchins (right) trapped spiders and other arthropods and found that coffee plantations in Hawaii that now host native trees and shrubs (left) have higher invertebrate diversity.

how the return of native flora affected arthropod diversity on these plantations.

He drove his sister-in-law's beat-up silver Toyota Tacoma to native farms and forests through a maze of branching roads and dirt tracks, maneuvering around gnarled roots, giant puddles, and solidified lava. At each of the handful of sampling sites he visited daily, Hutchins gathered leaf litter into a funnel with a lamp over it, which dried out the soil and caused arthropods to fall out into a collection tube. He erected large, tentlike structures to catch flying insects such as moths and flies, used brightly colored bowls to attract pollinators such as bees, and whacked trees with a stick to dislodge arthropods including beetles and thumbnail-size long-legged spiders, which he'd race to suck through a tube and trap before they scurried away. He'd do this from sunrise to sunset, interrupted only by afternoon rain showers that forced him to take shelter.

At first, incorporating Indigenous data sovereignty wasn't on Hutchins's radar. But 2 years into his Ph.D., around when he started his fieldwork, Hutchins attended an Indigenous data science workshop that inspired him to do more to build trust with the Indigenous farmers and seek their support for his project.

One way to build trust was to explore how his project could benefit the farmers. After gathering his data, he presented each farmer with a list of arthropod species identified on their farms, including pests—data more useful to them than to his project. “You need to engage in those conversations ... that can have meaning to people on the ground,” Hutchins says.

Hutchins was also aware that the arthropod samples he was collecting count as Indigenous data because they come from Native lands and have cultural significance, just like sacred objects and traditional knowledge. So he requested the farmers' consent for any data he shared in his paper, and redacted the names and locations of arthropod species and their sequencing data to keep culturally sensitive information from outsiders and reduce unauthorized visits to the farms.

Hutchins's approach to data sharing is a great example of Indigenous data sovereignty, says Nici Pfeiffer, chief product officer at the nonprofit Center for Open Science (COS), which works to sustain open research practices in science. Even though COS advocates for openness in the exchange of information, its data-sharing platform can accommodate

enous communities and researchers around the world to assert their data rights even as data repositories, scientific publishers, and federal and private funding agencies try to make data more transparent and widely accessible. People in the Indigenous community say that movement ignores principles developed to protect Indigenous people in the face of historical injustices. They also say open data initiatives don't do enough to ensure that communities see the benefits when their data are used, for example to develop new medical treatments. “Too often Indigenous Peoples have been told ... they need to contribute their DNA so that it benefits humankind, and unfortunately Indigenous Peoples are the last to benefit,” says Krystal Tsosie, a Diné/Navajo Nation citizen and geneticist-bioethicist at Arizona State University (ASU).

Yet change is coming. Many key players in the world of science—from the U.S. National Institutes of Health (NIH) and international journals to the governments of many countries—are rethinking their funding and data sharing structures to preserve sovereignty of Indigenous data as they get increasingly shared. Guiding the movement is a set

of principles developed by Indigenous scholars, called CARE (for Collective benefit, the Authority of Indigenous Peoples to control their data, the Responsibility to nurture respectful relationships with Native communities, and Indigenous Ethics in data sharing), developed to complement a push in some scientific circles for data sharing, openness, and collaboration. Indigenous researchers are also developing their own systems to curate data from their people and lands. “If that means upending the model so that Indigenous Peoples have proper stewardship and ownership of their data, then that's a data future that I'm optimistic for,” Tsosie says.

IN 2021, Hutchins spent many days on the west coast of the island of Hawai'i among dense forests and coffee plantations carved out of Native Hawaiian lands by American colonizers. The plantations originally grew only coffee, but farmers—many of them now Native Hawaiians—have recently started to restore native biodiversity, growing coffee among lush undergrowth in the shade of massive native fruit trees, including 'Ulu or breadfruit. For his Ph.D. project, Hutchins wanted to see

CARE principles, Pfeiffer says. In Hutchins's case, other researchers could discover his arthropods project and many details about it, but data Hutchins did not want discoverable, including species names and locations, could be restricted. "Data should be as open as possible, and as closed as necessary," she says, especially because "the risk is that [Indigenous data] is being misused or mishandled in a way that doesn't protect its sovereignty."

IT'S THIS HISTORY OF unethical use of Indigenous data that Hutchins and others are trying to change. Infamous examples include the misuse of blood samples collected by a University of British Columbia researcher from the Nuu-chah-nulth people in Canada for a study of rheumatoid arthritis. When that researcher left to go to a different institution, he took the specimens with him and also shared them with other researchers for unrelated studies—including on migration and retroviruses—without the knowledge or explicit consent of the Nuu-chah-nulth people. In a different instance, misuse of DNA samples collected from the Havasupai Tribe by ASU researchers led the Havasupai to discontinue all research and banish ASU researchers and employees from its reservation. The Havasupai filed a lawsuit against ASU researchers and the Arizona Board of Regents that ended in 2010 with \$700,000 in compensation to the tribe and the return of its citizens' DNA.

For many Indigenous community members who remember examples of such data misuse, the solution is to decide which data are shared and whom they are shared with. "It's not a free-for-all," says Rosie Alegado, an oceanographer at the University of Hawaii and a Native Hawaiian, adding, "'Accessible' means somebody can access it but you are still allowed to regulate it." To scientists who might feel strongly about unrestricted access to data, "You should check your ethics," she says. "That kind of thinking assumes best intentions, but [is] a little bit naïve, and within that naïveté is where bad actors can slip in."

Even when Indigenous communities have specifically consented to having their data used, such as for large-scale human genome sequencing efforts, they seldom benefit from the resulting scientific breakthroughs. "When we think about CRISPR and precision medicine, we don't have the system set up anywhere for Indigenous Peoples to benefit from those innovations," says Stephanie Russo Carroll, a data researcher at the University of Arizona who is Dene/Ahtna, a citizen of the Native village of Kluti-Kaah in Alaska.

In 2019, Carroll, fellow researcher Maui Hudson—affiliated with the Iwi of Te Whakatōhea, Ngāruahine, and Te

Māhurehure in Aotearoa (the increasingly used Māori name for New Zealand)—and their colleagues at the Global Indigenous Data Alliance established the CARE principles. In the 5 years since, these principles, which draw on existing ones among many Indigenous Peoples around the world as well as on the United Nations's Declaration on the Rights of Indigenous Peoples, have become foundational to the increased acceptance of Indigenous authority over their own data.

"The CARE principles have helped provide a roadmap for people to engage with these ideas," says Matthew Anderson, a geneticist at the University of Wisconsin–Madison who is of Eastern Band of Cherokee Indians descent. When he visited NIH in July to talk about a new Tribal Data Repository for COVID-19 data, the CARE principles made the concept of Indigenous data sovereignty easier to explain, he says. Hutchins says

"'Accessible' means somebody can access it but you are still allowed to regulate it."

Rosie Alegado
University of Hawaii

he also cites them when he raises the idea of Indigenous data sovereignty with non-Indigenous researchers. There may be a learning curve, but he hopes that, just as scientists came to embrace earlier ethical innovations like informed consent, Indigenous control of data will win broad acceptance.

WHEN HUTCHINS SEQUENCED the DNA of the arthropods in his collection, he learned that coffee plantations with Polynesian crops have an arthropod community more diverse than in monoculture coffee plantations. He was also reminded of the strong link between his work and Native Hawaiian culture. Native Hawaiians consider many of the arthropods he cataloged to be kin, and they are also protected by a 2003 Native Hawaiian declaration that asserts intellectual property and other rights over Native resources. Hutchins says he grew up hearing stories about native tree snails singing, and his paper includes examples, such as ants and grasshoppers, in the Kumulipo, a 2102-line creation chant describing the origin of the universe and various plants and animals.

Such context could enhance Western science, says Jane Anderson, an anthropologist at New York University, by providing otherwise overlooked scientific or conservation knowledge. "It's pretty amazing when you've got a community who has songs that have been sung for over a hundred years, naming

exactly the locations where the best museums are," she says. In 2010, Jane Anderson co-founded Local Contexts, which links data and specimens to unique digital tags that acknowledge Indigenous sources of knowledge and Indigenous data sovereignty. Museum collections, the Library of Congress, and scientific journals, databases, and projects increasingly use Local Contexts to credit Indigenous knowledge and provenance.

Hutchins, for example, created "biocultural" and "traditional knowledge" notices on the Local Contexts hub to indicate that his work on arthropods was subject to Indigenous data rights. Similarly, a University of Maine project to collect environmental DNA (eDNA)—shed by various organisms into the soil, water, or air—added a Local Contexts notice to indicate its commitment to working with the four Wabanaki Nations from whose lands they collect samples. Anytime the researchers collect eDNA samples from lakes, rivers, and the ocean to study shifts in community composition, Tribal representatives are automatically notified. They review the data to attach Local Contexts "labels" specifying their Indigenous provenance and indicating whether consent was given for their collection. Labels also specify how the data can be used: for research, outreach, commercial, or noncommercial purposes.

Using such notices is "maybe the first step of engaging in Indigenous data sovereignty," Hutchins says, although he emphasizes they're no substitute for jointly developing scientific projects and relationships with Indigenous communities.

WHEN IT CAME TIME for Hutchins to publish his research, he sought out a journal that he hoped would keep data sovereignty in mind, especially as some Indigenous researchers worry that storing information in university or journal databases means losing control of it. He ultimately chose *Molecular Ecology Resources*, whose special issue on Indigenous contributions to the field of molecular ecology included Indigenous editors.

To retain control of data, the journal allowed Hutchins to store his information in the Indigenous owned-and-operated Native BioData Consortium (NBDC) in Eagle Butte, South Dakota, nestled within the Cheyenne River Reservation of the Cheyenne River Sioux Tribe.

Data in NBDC are only available if researchers contact the nonprofit and get the permission of the specific Indigenous community whose data they are seeking. "We are a safe harbor," says Joseph Yracheta, Pūrepecha (Mexican Indigenous) and co-founder, along with Tsosie, of NBDC. Other researchers, like Hutchins, increasingly want



A coffee farmer (center left in green shirt) in Hawaii shows Leke Hutchins and his team around a plantation.

to store data there to ensure they are held on tribally sovereign lands under the corresponding legal frameworks.

NBDC is also where Matthew Anderson is storing data and samples for his project analyzing the general gut microbiome in the Lakota. “We’re not depositing any of this microbial information into any NIH database, despite ... policies and requirements around open data,” he says. He adds that several publishers have also allowed a statement in his papers telling researchers who want his data to request them from NBDC and the Cheyenne River Sioux Tribe. The goal is to receive collective consent from Tribes and their representatives rather than from individuals. “That really moves the power dynamic away from the [principal investigator] and towards the community,” he says.

NIH is beginning to recognize this shift. A new policy in the works may require researchers to get permission from Tribes to access their data in NIH data repositories, says Karina Walters, member of the Choctaw Nation of Oklahoma and director of NIH’s Tribal Health Research Office. She expects the draft policy to be ready for public vetting by early next year, with the final policy in place by September 2025. The policy may also require grant proposals to articulate how the project will build relationships with and benefit Tribal communities whose data are being sought. “The bottom line is that they will have to work closely with Tribes to find out what [they] want,” she says.

Indigenous scholars are also talking to other funders about changing their policies, noting that funders have powerful leverage over scientists. “If you can’t get the money, then you’re going to change the way you operate,” Matthew Anderson

says. CARE principles have shown up in calls for proposals by NSF, the U.S. Geological Survey, and private funding agencies. NSF has also started to require any proposals that could impact a Tribal Nation’s resources or interests to submit prior written approval from the Tribal Nation’s official representative (*Science*, 6 September, p. 1034).

Open science proponents are increasingly starting to acknowledge that Indigenous data need to be treated differently. COS’s Pfeiffer plans to work with Indigenous communities to implement data access with more options for sharing, consent, and licensing mechanisms for publicly shared data—rather than an all-or-nothing approach. And ESA’s open research policy now makes an exception for data, like Hutchins’s, sampled from Indigenous territory.

Major scientific publishers are also changing their policies to accommodate Indigenous sovereignty. International journals, including *Science* and those within Springer Nature, acknowledge the importance of the CARE principles. Springer Nature is considering how to implement them “in line with our commitment to our authors and supporting the ethical and sustainable sharing of information,” a spokesperson said.

But adoption of these principles across journals has been slow, says Mia Ricci, the American Geophysical Union’s director of publications operations. “I think a lot of people want to do this, they just don’t know where to get started,” Ricci says. Her organization has been working with Carroll’s team as well as with many publishers on drafting broad guidelines for implementing Indigenous data sovereignty and the CARE principles, work that they plan to present at an NSF-funded workshop in February 2025.

In April, hundreds of scholars, researchers, and Tribal leaders from all across the world gathered for the first ever Indigenous Data Sovereignty Summit in Tucson, Arizona, with the goal of continuing to help individual Tribal nations develop their own laws around data and inspire the U.S. to establish national standards for handling such data, similar to existing policies in Canada, Australia, and Aotearoa. “Ideally there comes a time when this becomes the status quo, when we’re no longer fighting to ... even recognize that Indigenous data sovereignty exists,” says Desi Small-Rodriguez, Northern Cheyenne and Chicana, a social demographer at UC Los Angeles and co-founder with Carroll of the U.S. Indigenous Data Sovereignty Network.

For now, many non-Indigenous researchers still haven’t heard about Indigenous data sovereignty principles, let alone how to implement them. And whereas open data policies are increasingly required by funders and publishers, Indigenous principles are still generally considered optional and up to individual researchers.

“Too often in science, we put the onus for ethical behavior on individuals, but an individual cannot ultimately behave ethically within an unethical framework,” Carroll says.

Despite the challenges, Hutchins is seeing the shift in real time. “How do you change the culture of a lab that has all these decades of doing things a certain way?” he wonders. Hutchins may not have the full answer, but he says, “I think things are starting to slowly turn in the right direction.” ■

Sandeep Ravindran is a science journalist near Washington, D.C. This story was supported by a grant from the Heising-Simons Foundation.