From Green Revolution to "Rescue" Indigeneity:

Using Digital Data Tools and Machine Learning Approaches to Protect Indigenous Knowledges and Biodiversity

KRYSTAL TSOSIE, PHD, MPH, MA

ARIZONA STATE UNIVERSITY

NATIVE BIODATA CONSORTIUM

@KSTSOSIE

Krystal Tsosie to be ASU's first Indigenous human geneticist

November 4, 2022

Krystal Tsosie (Diné/Navajo Nation) is an advocate for Indigenous genomic and data sovereignty. She is a co-founder of the first U.S. Indigenous-led biobank, a 501(c)3 nonprofit research institution called the Native BioData Consortium. Her current research at Arizona State University centers on ethical engagement with Indigenous communities to ensure Indigenous peoples equitably benefit from precision health and genomic medicine.

Tsosie's research and educational endeavors have received international media attention in such outlets as The New York Times, PBS NOVA, Washington Post, NPR, The Atlantic, Forbes and The Boston Globe, among others.

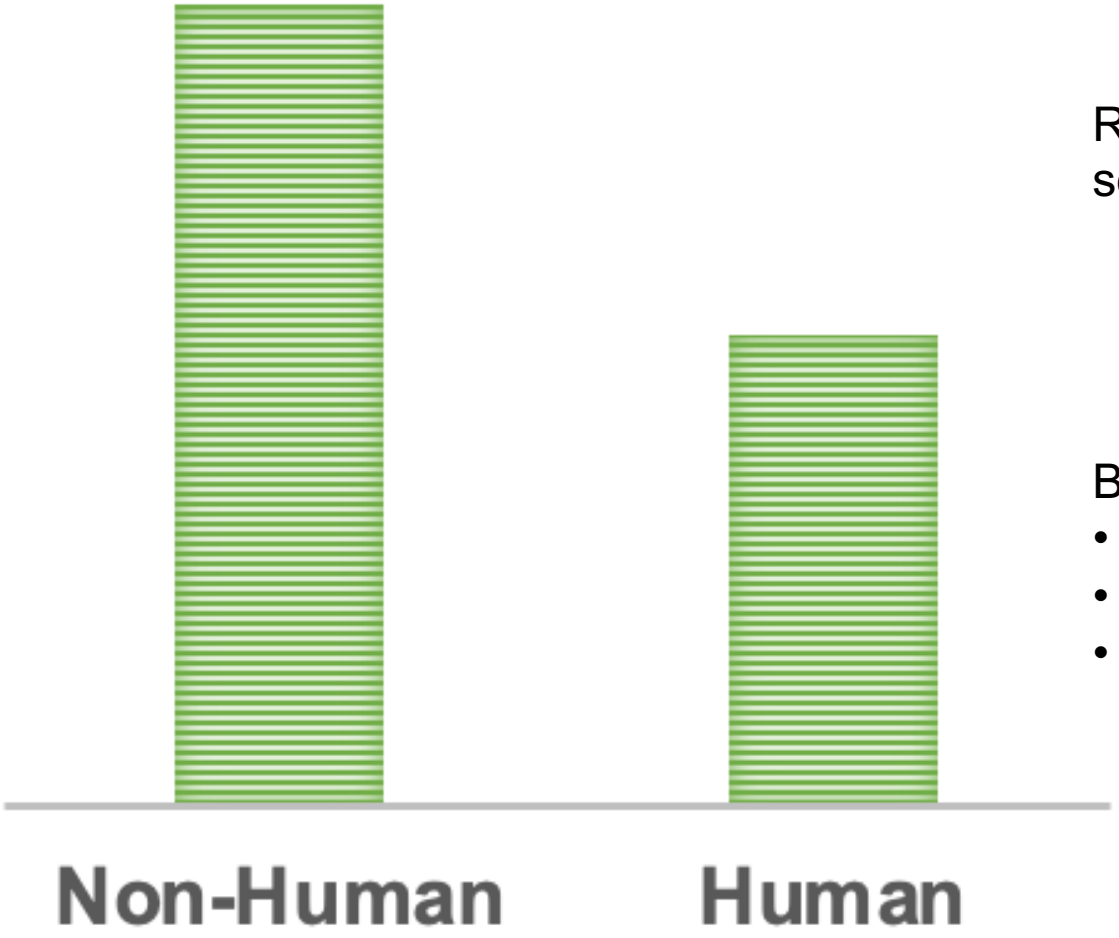
Krystal Tsosie, an ASU Presidential Postdoctoral Fellow at ASU, will become an assistant professor in the School of

She is a current ASU Presidential Postdoctoral Fellow, and starting Jan. 1, 2023, she will be an assistant professor in the School of Life Sciences at The College of Liberal Arts and Sciences and the first Indigenous human geneticist at ASU.

"I'm excited to come back to ASU and serve as an advocate for Indigenous communities. I want to bring all of these skill sets related to health inequities and genetic epidemiology back to the communities that I grew up with," said Tsosie.

She is also one of two scholars selected as a 2022–2023

Why Focus on Eukaryotic Non-Humans as a Human Geneticist?



Research on “non-humans” is less regulated, so the pace of research is moving a lot faster.

Both spaces grappling with similar Qs:

- How to ethically collect data?
- How to manage data access and sharing?
- How to advance sequencing technology?

NATURAL HISTORY MUSEUM

SCIENCE NEWS

Are natural history museums inherently racist?

By Josh Davis

First published 16 July 2019

Much of biology is rooted in the colonial act of "discovering" natural specimens and **collecting them in museums and biocollections** for the sake of science.

Often these collections claim credit for medicinal and plant knowledges from enslaved Africans and Indigenous peoples but attribute that science to naturalists instead.

"Gilded Canopy" – mural depicting plants that fueled the British Colonies

Das, S. & Lowe, M. (2018). Journal of Natural Science Collections

Pre-Settler-Colonial State

- Estimated 30,000 edible plant species worldwide
- Indigenous agrisystems considered “**backwards**” or “**primitive**”
- Many cash crops and medicines from colonized lands and knowledges

Green Revolution

- Only 30 plant species constituting majority of diets
- 60% biodiversity loss
- Contributing to 1/3 of greenhouse gas emissions
- Degrading local ecosystems
- Inequities: power, wealth concentrated by corporations

“Rescue” Indigeneity

- Indigenous knowledges “**key**” to sustainability
- Seeds and varietals stewarded by Indigenous peoples deposited in global seedbanks and used from museums
- Gene editing approaches on the horizon to reintroduce biodiversity

In order to advance genomic science...

we need more data that is freely accessible.

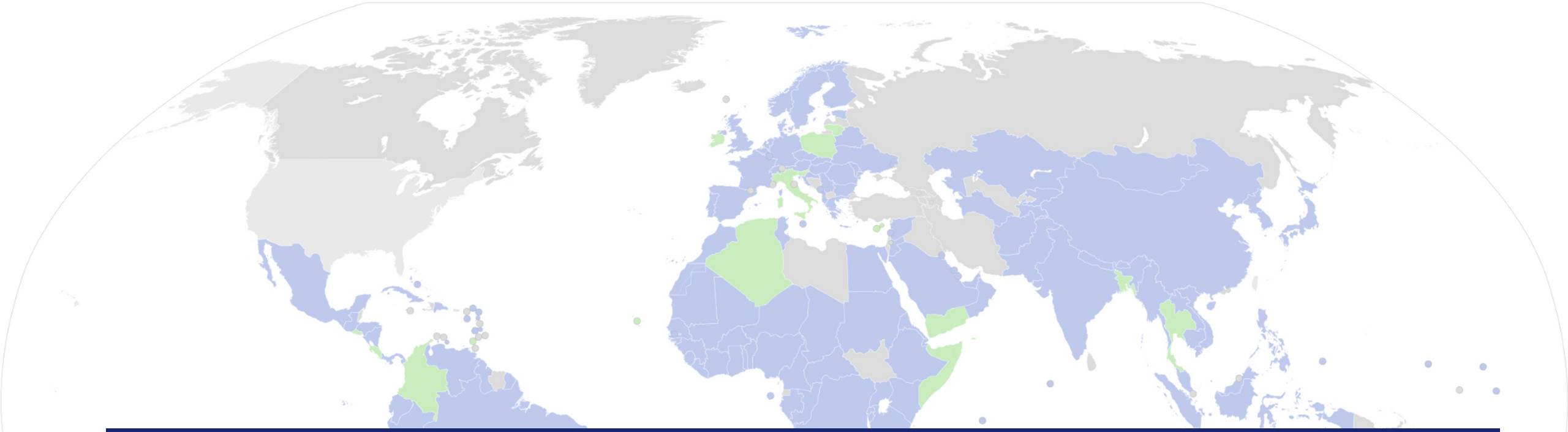
However, recent and past unethical research conduct involving the collection of data from Indigenous communities has strained trust relationships, resulting in **Tribal Nation policies that restrict open data sharing**


Comprising less than 5% of the world's population, Indigenous people protect 80% of global biodiversity.
The next genomic ”discoveries” may co-opt Indigenous knowledges or disenfranchise Indigenous peoples, who are often last to benefit and are least protected from intellectual property claims.

INDIGENOUS GENOMIC DATA SOVEREIGNTY

The right of Indigenous nations and people to exercise autonomy to protect their interests related to genomic data.

This “sovereignty” is intrinsically-defined, and not colonially-defined.



The **Nagoya Protocol,** a legal framework under the Convention on Biological Diversity (CBD), formalizes fair and equitable sharing of benefits arising from biological diversity.

It encompasses biological samples and associated Indigenous knowledge, with equitable return of benefits to those providing samples.

How can Indigenous peoples operationalize the Nagoya Protocol?

Guiding Principles for Data Governance and Stewardship

Wilkinson et al. *Scientific Data* 2016. **The FAIR Guiding Principles for scientific data management and stewardship**

Carroll et al. The Global Indigenous Data Alliance. 2019. **CARE Principles for Indigenous Data Governance**


Native BioData
consortium

Native BioData Consortium (NBDC)

- The first Indigenous-led biological and data repository ("biobank") within tribal jurisdiction
- A 501(c)3 non-profit corporation organized under tribal and Federal law
- Board members from Cheyenne River Sioux Tribe and other Indigenous nations
- Research partnerships since 2009
- Founded and incorporated in 2018

Native BioData consortium

DIGITAL TOOLS TO PROTECT INDIGENOUS GENOMIC DATA SOVEREIGNTY

- ✓ **Traditional Knowledge (TK) Labels.** Digital markers that define attribution, access, and use rights for Indigenous cultural heritage
- ✓ **Biocultural (BC) Labels.** Digital markers for provenance, transparency and integrity in research engagements related to community expectations and consent for use of collections and data.

- ✓ **Dynamic Consent Portal.** In Indigenous-led data repository to house Tribally-consented genomic sequence data and manage access and attribution.

- ✓ **Blockchain.** A distributed ledger system that tracks sharing via transactions, can fine-tune user access, attribute provenance, and facilitate data governance.
- ✓ **Federated learning.** To facilitate secure and community-consented data sharing.

Local Contexts

Native BioData consortium

Cell

Log in

COMMENTARY | VOLUME 185, ISSUE 15, P2626-2631, JULY 21, 2022

Establishing a blockchain-enabled Indigenous data sovereignty framework for genomic data

Tim K. Mackey • Alec J. Calac • B S Chenna Keshava • Joseph Yracheta • Krystal S. Tsosie • Keolu Fox

DOI: https://doi.org/10.1016/j.cell.2022.06.030 •

Comment

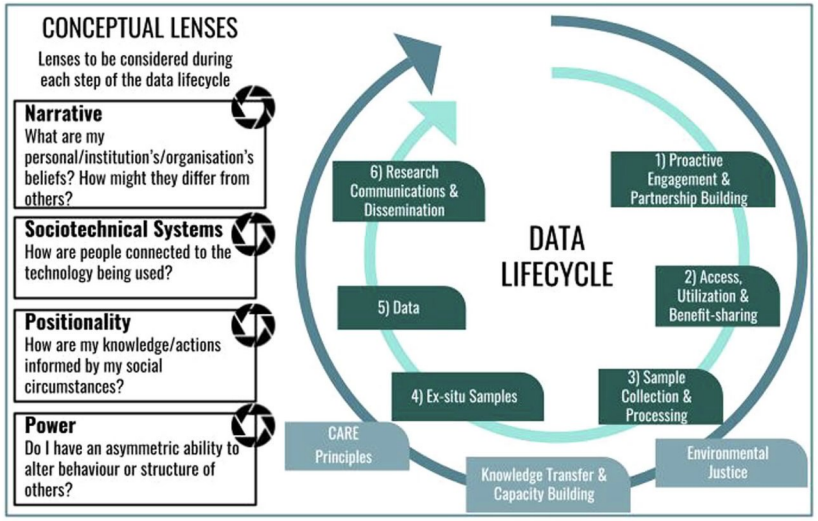
https://doi.org/10.1038/s42256-022-00551-y

Federated learning and Indigenous genomic data sovereignty

Nima Boscarino, Reed A. Cartwright, Keolu Fox and Krystal S. Tsosie

Indigenous peoples are under-represented in genomic datasets, which can lead to limited accuracy and utility of machine learning models in precision health. While open data sharing undermines rights of Indigenous communities to govern data decisions, federated learning may facilitate secure and community-consented data sharing.

data decision-making authority outside of Indigenous data governance structures. Unlike for other subsets of individuals who may be included in genomic research, the United Nations Declaration on the Rights of Indigenous Peoples (UNDRIP) recognizes and respects the self-determination and self-governance of Indigenous peoples[7]. This includes the right to "maintain, control, protect and develop [...] their technologies [...], including human and genetic resources," which is exemplified by calls from global Indigenous peoples for the cessation of large-scale diversity genome projects in their communities[8]. Despite efforts over the past 20 years to increase the inclusion of under-represented populations in datasets, as of 2019 Indigenous



CARE Principles applied to biodiversity genomics researchers

Collective Benefit: Research with IPLCs should function, support, and promote IPLC use and reuse of data and facilitate exploration, articulation, and application to enhance production of valued outcomes for IPLCs.

Authority: IPLCs have the right to control research conducted in partnership with IPLCs, and on IPLC lands. Indigenous sovereignty and data sovereignty must be embedded into projects, continued throughout, and maintained after completion.

Responsibility: Conducting culturally safe research in partnership with IPLCs requires 1) cultural, spiritual and epistemological competency; 2) an understanding of the extractive, deficit-based, exploitive research that continues to impact IPLCs; and 3) an awareness that both 1 & 2 could impact relationships between IPLCs and the research enterprise.

Ethics: To maximize benefit, minimize harm, and promote justice many IPLCs have laws, regulations, or policy on the collection and sharing of genetic resources, associated sequencing information, and aTK. Understanding and respecting these laws, processes, and policies is critical when planning and conducting ethical research in partnership with IPLCs.

Pillars of Environmental Justice

Procedure
Perception of fairness is impacted by the quality of the experience throughout the research & not only the outcomes of those experiences.

Recognition
Acknowledgement of the political, legal & knowledge systems, including any harms of research.

Distribution
Fairness of how risks and benefits are shared between across the research project.

npj | biodiversity

www.nature.com/npjbiodivers

REVIEW ARTICLE OPEN

Check for updates

Indigenous peoples and local communities as partners in the sequencing of global eukaryotic biodiversity

Ann. M. Mc Cartney[1✉], M. A. Head[2], K. S. Tsosie[3,4], B. Sterner[4], J. R. Glass[5], S. Paez[6], J. Geary[7] and M. Hudson[2]

INDIGIDATA

Indigenous Data Science Education

IndigiData.org/apply/

Now Offering 2 Workshops This Summer
Apply By Feb 28, 2023

#1

INDIGIDATA MN
JUNE 11 - 17, 2023

Gene Editing, Modification, and Food Sovereignty

Discuss Indigenous knowledges related to genetic modification, traditional foodways, and CRISPR gene editing

#2

INDIGIDATA WA
JUL 30 – AUG 5, 2023

Waterways, Fisheries, and Metabarcoding

Indigenous fish management, and environmental conservation using qPCR/metabarcoding

IndigiData.org

INDIGIDATA

Indigenous Data Science Education

WORKSHOP INFO CURRICULUM FACULTY PARTNERS NATIVE BIODATA CONSORTIUM

empowering the next generation of Indigenous data scientists

INDIGIDATA INFO

Native BioData consortium

krystal.tsosie@asu.edu

www.nativebio.org

New Tsosie Lab for Indigenous Genomic Justice