

Commentary

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

Tim K. Mackey,^{1,2,3,4} Alec J. Calac,^{3,5,6} B S Chenna Keshava,⁷ Joseph Yracheta,⁸ Krystal S. Tsosie,⁸ and Keolu Fox^{1,8,9,*}

¹Global Health Program, Department of Anthropology, University of California, San Diego, San Diego, CA, USA

²S-3 Research LLC, San Diego, CA, USA

³Global Health Policy and Data Institute, San Diego, CA, USA

⁴BlockLAB, San Diego Supercomputer Center, San Diego, CA, USA

⁵University of California, San Diego, School of Medicine, San Diego, CA, USA

⁶University of California, San Diego, Herbert Wertheim School of Public Health and Human Longevity Science, San Diego, CA, USA

⁷University of California, San Diego, Department of Computer Science and Engineering, San Diego, CA, USA

⁸Native BioData Consortium, Eagle Butte, SD, USA

⁹Indigenous Futures Institute, Division of Design and Innovation, University of California, San Diego, San Diego, CA, USA

*Correspondence: pkfox@ucsd.edu

<https://doi.org/10.1016/j.cell.2022.06.030>

Technological advances have enabled the rapid generation of health and genomic data, though rarely do these technologies account for the values and priorities of marginalized communities. In this commentary, we conceptualize a blockchain genomics data framework built out of the concept of Indigenous Data Sovereignty.

Background

The movement toward precision medicine has spurred a growing debate regarding the access, management, and ownership of individual health data. Complicating the matter, despite an explosion in the volume and diversity of individual health information collected, data access largely remains outside the control of patients or communities who could benefit from their research or commercial application. Equally, the digitization of social and genetic health data outside of the clinical setting has introduced new challenges that require a reconceptualization of data privacy considerations (Bari and O'Neil, 2019). Further, discoveries derived from these data rarely provide compensation or immediate benefit to participants and may actually reinforce stereotypes and biases against marginalized communities, such as identifying "risk" alleles for substance use, rather than prioritizing community-based interventions (Kinchin et al., 2017; Melroy-Greif et al., 2016). For these communities, and specifically Indigenous peoples, challenges are compounded with additional concerns about secondary use of data, unique privacy

considerations, and even more limited opportunities for benefits sharing.

Genomic data from Indigenous peoples in particular have been the target of researchers interested in advancing our understanding of various human diseases and genetic disorders and conditions, often used without free, prior, and informed consent (Tsosie et al., 2021). Indigenous populations, historically underrepresented in research, have a relatively low degree of European admixture and offer unique insights into genetic variants of interest. In 2021, the NIH All of Us Research Program (a research program focused on building a diverse health history database involving more than one million people) published guidelines for responsible tribal engagement, ensuring that health information specific to Indigenous Americans is protected at all stages of the research process (National Institutes of Health, 2021). This announcement was met with skepticism given concerns about tribal involvement in genetic research and no explicit focus on returning material benefit to tribes (Fox, 2020).

We recognize that individual participation in research can have sociopolitical implications for sovereign Indigenous na-

tions. However, there are several challenges to encourage Indigenous nations' participation in genetic research and operationalize Indigenous Data Sovereignty (IDS) (see Box 1 for definition). First, researchers and funding structures may not recognize the authority of Indigenous nations to convene their own Institutional Review Board (IRB) or agree to provide Indigenous nations full data access and ownership (Harding et al., 2012; Around Him et al., 2019). Second, existing data systems often misclassify Indigenous peoples as members of other racial and ethnic groups, leading to inaccurate data and failure to generate the evidence needed for advocacy and policy making (Yellow Horse and Huyser, 2021). Finally, Indigenous nations must often navigate complex jurisdictional environments in trying to exert control over their health data (Walker et al., 2018).

These issues warrant the development of data systems that operationalize IDS and equity, ensuring that each Indigenous nation can control access to their members' sensitive health information and prioritize research that is linked to community priorities. IDS anticipates the purposeful inclusion of Indigenous peoples

Box 1. What is Indigenous data sovereignty?

Indigenous data sovereignty (IDS) is defined as the **right of an Indigenous nation to govern the collection, ownership, and application of data generated by its members**. However, IDS implementation through innovative technologies such as blockchain has not actively been explored in the literature.

in the governance and design of systems that securely manage their biological samples and health data in a transparent manner, while also acknowledging the cultural values and interests unique to each individual Indigenous nation.

While there are digital tools such as digital biobanks and citizen science platforms supporting the generation and commercial sharing of health data, they largely do not enable Indigenous nations to exercise direct control over their data due to a lack of purposeful system design, institutional support, and authority over these proprietary systems (Tengo et al., 2021). Therefore, these public data environments may indirectly enable the continued exploitation of Indigenous peoples' data in violation of Indigenous nations' rules and regulations. This technology gap presents an opportunity to bridge innovative technologies and Indigenous knowledge systems, ensuring that new data management technologies are co-governed with Indigenous nations.

Blockchain, genomics data, and Indigenous populations

Blockchain technologies, which are a form of distributed ledger technology popularized by cryptocurrencies such as Bitcoin,

enable data provenance, increased transparency, and enhanced trust within a distributed network (see Box 2 Glossary). Indigenous-led non-profit organizations, such as the First Nations Technology Council in Canada, are exploring the use of cryptocurrencies to minimize government involvement in Indigenous affairs and to develop resource management platforms. Specific to healthcare, blockchain is being explored as a solution to better manage data, including growing commercial interest in genomic data applications (Tandon et al., 2020). There is a general tendency to use blockchain technology to decentralize and democratize the storage and use of health data (Miyachi and Mackey, 2021). As a blockchain represents a permanent, near-immutable ledger of transactions, generally, blockchains can be used to keep track and help mediate access to health and genomic records or used to enable sharing of data attributed to a validated digital identity. However, the merits of blockchain's use and compatibility with IDS or Indigenous genomic data have not been examined.

Importantly, Indigenous peoples' participation in genomics research may be limited by their respective Indigenous nations. For example, the Navajo nation, one of the largest Indigenous nations in

the US, has had an active moratorium on human genetics research for nearly two decades in response to the misuse of genetic samples. Further, more fundamental infrastructural challenges may hamper digital transformation initiatives or the adoption of new technologies. Specifically, information technology systems for social and healthcare agencies serving Indigenous peoples are more likely to be antiquated or under-resourced. Complicating this, many reservations (areas in the US designated for Indigenous peoples' use and housing) lack internet access and electricity, creating a digital divide that furthers inequity. Hence, genomic biobanks, big data initiatives, and new technologies such as blockchain will require extensive infrastructure development and resourcing.

Though these challenges need to be addressed through community-centered technology resourcing, implementation, and evaluation, it is also necessary to conduct a more purposeful assessment of whether technologies such as blockchain actually align with IDS principles and how technology features can be adapted to specific community needs. While innovations in digital health now focus on "patient-centered" approaches putting the patient and their values at the center of a collaborative design approach, what we seek in this assessment is to develop the early foundations of an IDS blockchain framework that is community centered and enables Indigenous peoples to engage in distributed sovereign data management.

Box 2. Glossary

Blockchain: Blockchain systems are primarily composed of a distributed ledger that records transactions and is shared and agreed upon by all parties as the sole record of transactions (with agreement on transactions established through a process known as a consensus mechanism); a cryptographic hash function (used to generate a value to cryptographically link series of "blocks" of data, ensuring their security and near immutability); and a series of nodes (e.g., computers in a peer-to-peer network) that make up the network that operates the blockchain. In a blockchain, each block also represents a collection of data about past transactions, with every new block containing the hash of the former block, thus creating a "blockchain" of timestamped data establishing the agreement, provenance, and finality of the history of a transaction or management of data between a network of users.

Consensus mechanism/protocol: Allows distributed systems to reach agreement on what is written to the blockchain.

Consortium blockchain: Generally, describes a blockchain controlled and governed by a group.

Decentralized autonomous organizations: A software-enabled organization with no central authority built and governed by smart contracts on a blockchain network.

Distributed applications (dApps): Digital applications or programs that exist and run on a blockchain or peer-to-peer network of computers.

Node: A copy of the ledger operated by a user on the blockchain.

Private blockchain: Generally, describes a blockchain controlled by one authority.

Proof of authority: Consensus mechanisms based on identity as a stake.

Proof of stake: Consensus mechanism where those with the largest holding of the network's currency validate new blocks.

Public blockchain: Generally describes a blockchain open to public participation and which has no central authority.

Token: Representation of a digital asset built on an existing blockchain.

The need for blockchain-based Indigenous data sovereignty

The rationale for why an IDS genomics blockchain needs to be co-created and governed by Indigenous nations is motivated by the recognition of these diverse groups as self-governing entities able to regulate their health and political affairs. For example, existing medical and public health efforts with American Indian and Alaska Native Tribes prioritize local control of funds and programs, allowing for community members and healthcare leaders to determine the best processes for design, implementation, and sustainment of community-directed healthcare programs supported by the Indian Health Service. Further, given past exploitation and the need to strengthen ownership and governance of Indigenous data directly by communities themselves, certain Indigenous nations have taken the lead in developing their own capacity for collecting and managing genomic data.

One prominent example is the Native BioData Consortium (NBDC) in Eagle Butte, South Dakota. The NBDC, a first-of-its-kind non-profit Indigenous-led biobank and research institute, ensures that advances in genetics research provide material or immaterial benefit to Indigenous nations and hosts skills training workshops for aspiring Indigenous data scientists. In 2021, Illumina, Inc. donated several high-speed sequencers to the NBDC, greatly increasing the capacity of the organization to conduct on-site sample collection and analysis. Crucially, the NBDC is a biobank led by Indigenous scientists and community members operating within the jurisdiction of tribal lands and which has its own dedicated infrastructural capacity. This “sovereign” architecture, which keeps biological samples and data within the community and also simultaneously builds Indigenous research capacity, better ensures equitable benefits sharing in alignment with shared community goals.

Recognizing community buy-in of these initiatives, new technologies, such as blockchain, should not seek to reinvent governance structures and existing community partnerships, but instead, enhance them by enabling self-governance of data systems. Hence, our conceptualization of an IDS blockchain framework begins by mapping its features to structures already established by the

NBDC, while also exploring how the technology can augment these efforts by enabling distributed governance, dynamic community consent, and creating an immutable record of research and commercial actions for those stakeholders who wish to participate in Indigenous genomic data discovery.

Conceptualization of an IDS blockchain framework

Here we describe the basic conceptual framework for a genomics IDS blockchain framework that adopts IDS principles already established through the NBDC. There are a few core principles of IDS that need to be integrated into the blockchain systems’ principal design architecture. First, the blockchain should incorporate individual- and community-level data ownership and access privileges, thereby enabling novel distributed community-based governance strategies that involve both the representatives of different Indigenous groups, but also the **Indigenous community members themselves**, as well as **limited participation of external non-Indigenous entities that seek to access data in the system for scientific purposes agreed upon by the community**. In this context, we adopt a “consortium”-based blockchain design approach (distinct from public and private blockchains), where participation in the blockchain involves multiple pre-vetted tribes and organizations with different authorities and levels of permission, but which inherently represents a decentralized network of participants that all agree to the shared governance principles of the network (see [Table 1](#)).

Second, the set of rules, roles, and responsibilities that govern the consortium blockchain will be designed and specifically adhere to the core principles of IDS that establish Indigenous peoples’ right to govern the collection, ownership, privacy, and application of their own data. This is distinctly different than other data management systems that are not indigenous specific as IDS will be purposefully coded into the design of the blockchain’s governance environment, including through the execution language of smart contracts (i.e., programs or code stored, called, and executed on a blockchain when predetermined conditions are met) used for data access and

sharing. Further, regulatory issues will also be made compatible with IDS and incorporate other applicable tribal IRB, ethics, and privacy requirements that are unique to specific Indigenous peoples.

Third, enabling community engagement and data management aspects of the IDS framework will use community consensus mechanisms that are designed to achieve agreement and validate transactions among the distributed nodes. There are multiple options for the implementation of consensus mechanisms, though our approach focuses on proof of authority (POA) and proof of stake (POS). In a POA-based consent mechanism, at least 51% (simple majority) of the authorized representatives or members of an Indigenous nation would need to accept a data request for successful access, otherwise, it will be denied. Alternatively, a modified POS consensus can be used, similar to use cases in Decentralized Finance dApps, with the stakeholders holding the largest volume of data having a higher stake or say in approving or rejecting a data request. Since we have not introduced tokens in the IDS framework, stake can be measured in terms of the data contributed and/or subject to a specific data access request. This would ensure that communities whose members’ data is being requested, retain the most “stake” in deciding whether said access should be granted. Consent mechanisms can be customized based on the proportion and type of data provided by each Indigenous nation with votes weighted based on the proportion of the genomic data records subject to an access request.

The overall goal of the IDS blockchain framework is to enable distributed community-mediated management of Indigenous genomic data, digital sequencing information, and associated metadata or state of data that can be useful for non-Indigenous researchers, biotechnology life science companies, and innovators in a privacy-preserving manner (see [Figure 1](#)). In order to do this, technical features of data management will focus on tribal- and community-level permissions, secure multi-factor consent and authorization, and utilizing an off-chain blockchain data management design in order to ensure that underlying genomic data is not shared directly on the blockchain ([Miyachi and Mackey, 2021](#)).

Table 1. Examples of roles and responsibilities of stakeholders on the IDS blockchain framework

| Stakeholder | Description of role | Responsibilities |
|---|---|---|
| Indigenous representative or organization | There are several design approaches that can be explored. Indigenous organizations can act as master nodes on the IDS framework to facilitate other events (e.g., governing voting events, execution of protocols, creation of smart contracts, and enforcing governance principles). Alternatively, they could also act as blockchain oracles and allow interface between smart contracts and on-chain and off-chain assets operated through decentralized applications (dApps). | Authorized Indigenous representatives or organizations can have authority delegated to them by individual tribal members or nations. The primary purpose for this delegation is to allow these organizations to administer the governance of the framework in a manner consistent with community values and Indigenous regulations, set the terms of smart contracts and consensus operated on the IDS framework, validate nodes to be added to the network, and allow querying, verification, and authentication of external data sources as needed. |
| Individual Indigenous member | Generally, individual Indigenous members act as “authority nodes” on the IDS framework, have the ability to access the network, can accept or reject transactions, can create and validate blocks, and have visibility to all transactions written to the blockchain. | Individual Indigenous members can have their digital identity validated through Indigenous eligibility verification processes (e.g., LIHEAP Clearinghouse) in order to gain privileges. Their primary responsibility is to exercise their authority to vote and reach consensus on data access requests that are then written to the blockchain. However, if necessary, this authority can also be delegated to master nodes (Indigenous organizations). |
| External non-Indigenous member parties | Non-Indigenous entities are parties that have read-only access to data written to the blockchain. They have no authority or voting rights, and only have the ability to query de-identified metadata in a privacy-preserving fashion that is written to the blockchain, with these queries all logged for transparency to the community. These members can be removed and added per consensus of the network and are vetted prior to joining the consortium. | Non-Indigenous entities primarily consist of researchers, life science and biotechnology companies, genomic companies, other biobanks, and other public and private entities that are engaged in Indigenous genomic data discovery. Their participation on the blockchain subjects them to the IDS governance rules, smart contract terms, and other protocols required and agreed upon by Indigenous nations, their IRBs, and other applicable Indigenous regulations. Data access is mediated through off-chain mechanisms with only the decision of the network to approve or reject a request for data written to the blockchain. |

This interplay between on-chain and off-chain data storage, processing, and management will be key features of the technical framework proposed. Specifically, hybrid off-chain blockchain systems can enable better scalability, reduce data storage requirements (important due to the large size of genomic data), allow for data querying, and enhance data privacy options (Miya-chi and Mackey, 2021).

Specifically, **only authorized nodes (pre-vetted organizations or individuals admitted to the consortium and who agree to the IDS governance principles)** will be able to make data queries. No

explicit individually identifiable information will be written on the blockchain to better ensure that the system preserves privacy. Instead, de-identified metadata (e.g., gender, age group, data type - array data, DNA or RNA sequence data, whole-genome sequencing, and possibly phenotypic data) may be stored on-chain and used to map off-chain data sources. This metadata can be queried for attributes of interest, wherein the data requests, smart contract access control, and the result of the query will be written to the blockchain in a transparent and immutable manner. Once a requestor identifies records of interest with

specific metadata attributes, their access approval will be included in the response that is written to on-chain storage. Upon a user having their query access approved, a user can then gain access to the underlying data from an off-chain data source. The off-chain data source can query the on-chain storage to validate that the user was approved to access certain data. Furthermore, the off-chain data storage could record user access to records as well.

Importantly, the location of an individual’s genomic data continues to reside external to the blockchain on the off-chain source, such as the tribal biobank servers

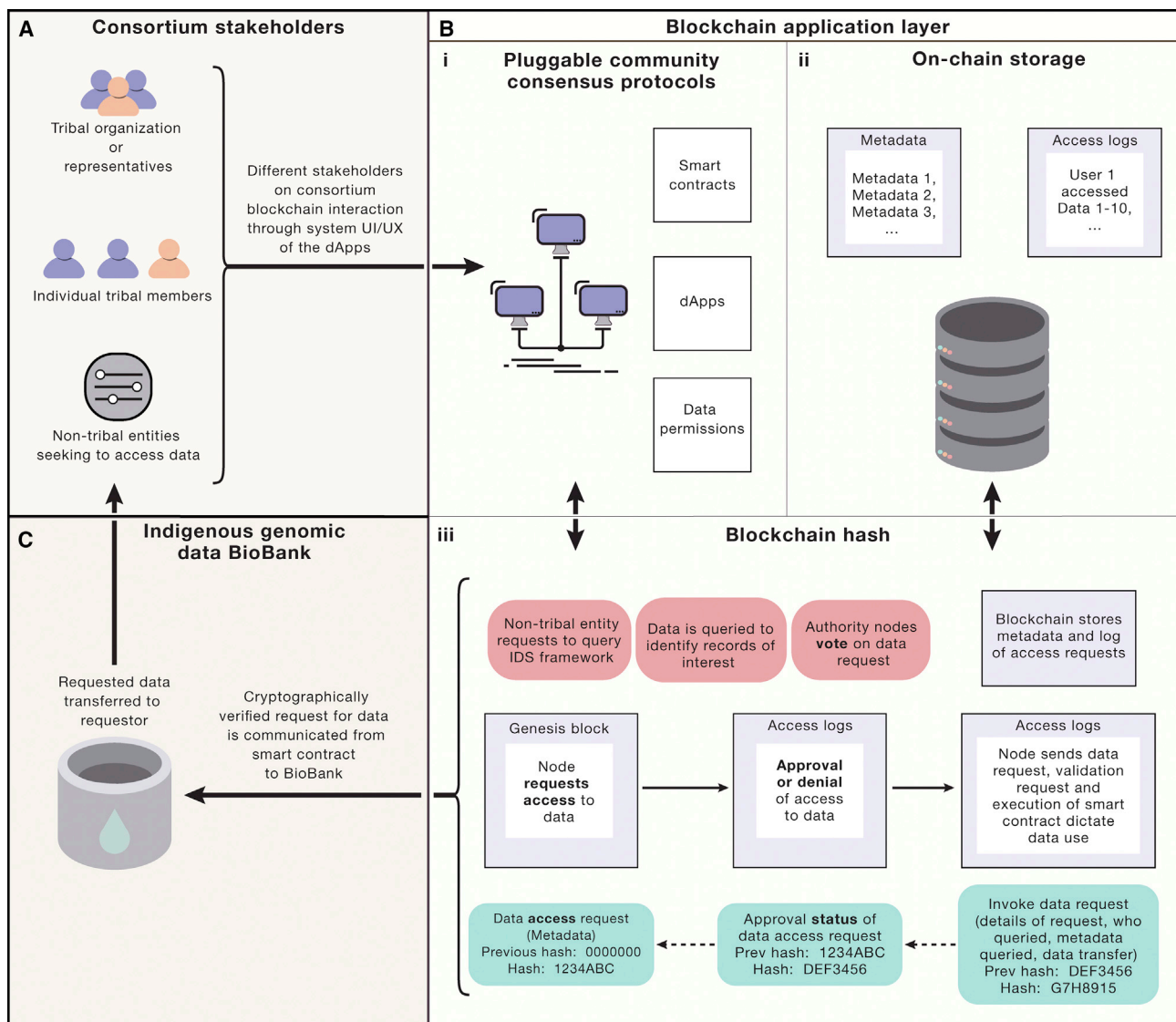


Figure 1. IDS blockchain framework summary

This figure describes a high-level architectural overview of the IDS blockchain framework. In the top-left corner (A), the different stakeholders who act as nodes on the blockchain interact with the blockchain via the smart contract user interface (UI). The blockchain is comprised of certain essential blockchain features including community-centered consensus protocols, smart contracts, permissions, and any necessary DApps (B-1). On-chain storage of data is de-identified and privacy preserving, with only a thin layer of metadata available for parties to query for purposes of identifying matching records that may be of interest for further genomic data discovery (B-2). The blockchain process described in (B-3) describes how blocks of data will be written to the chain based on a user querying metadata, adjudication by the network to allow or deny the request, writing the final result of said request to the blockchain, and executing the smart contract that outlines the terms and requirements of data release. Finally, upon receiving validated authorization as written to the blockchain and evidenced by consensus, the biobank as a node to the network can view and use this information to initiate transfers of requested data directly to the intended recipient via other off-the-chain mechanisms with the grant-of-access recorded on the blockchain (C)

(e.g., the NBDC biobank), with access to records mediated off-chain by cryptographic (SHA256) hash and pointers to the location of these records (e.g., genomics data locators residing on biobank servers off-chain). Further, with the data request successfully approved by the community through consensus, a smart contract can automate the execution of

an agreement to mediate access to the genomic data off-chain to the requestor through the exchange of secure public-key cryptographic encryption, similar to the model used for MIT's MedRec platform that manages medical data using smart contracts and decentralized content management (Ekblaw and Azaria, 2016).

Future considerations

While blockchain has the potential to lead to the “democratization” of data through distributed governance, there is a general lack of support for incorporating the values and priorities of marginalized communities into the development and management of these technologies. In this commentary, we identified key strategies

needed to align blockchain technology to IDS principles, with a focus on conceptualizing technology that augments organizational capacity for privacy and data autonomy and does not reinvent existing governance and community partnership structures. This technology framework has the potential to establish an innovative distributed governance structure that could functionally enable the management, sharing, and use of Indigenous genomic data consistent with the values and priorities of the sovereign governments from which they originate.

One caveat: we acknowledge that this is an early version of a framework co-ideated with NBDC board member co-authors. To improve our process we plan to prioritize an assortment of community consensus-building tools to seek broader community input, engagement, co-creation, and consent from involved Indigenous leaders and their constituencies through community-based participatory research approaches, particularly around establishing shared understanding of concepts related to privacy and security (Claw et al., 2018). Further, evaluation of what blockchain platforms might represent the optimal environment to build the IDS framework on will be important considerations. For example, Hyperledger Fabric's permissioned blockchain environment may offer greater scalability, performance, and security of data while also retaining the ability for rich querying. Ethereum, a popular decentralized blockchain software platform, offers a robust smart contract execution environment and the ability to stand up transparent community-led decentralized autonomous organizations (DAOs) governed by smart contracts.

These considerations should not delay continued community consultation, further modification and iteration, and co-ideation with Indigenous peoples of how this framework might serve as a starting point to develop a data governance environment that better meets the needs of specific Indigenous nations. Additionally, opportunities for empowering data decentralization as a means for sustainable economic benefit generation can also be considered through introduction of IDS-specific cryptocurrencies/tokens or non-fungible tokens (NFTs) that operate for the sole purposes of ensuring

that benefits from data sharing come back to community members and are tracked in a way transparent to measuring the impact of these benefits. Finally, promoting climate resilience remains a priority for Indigenous peoples. Given the heavy carbon footprint of some blockchain-based data governance ecosystems, we hope that engineers, data scientists, and traditional knowledge keepers can co-design natural and built environments that promote and empower environmental sustainability, invest in natural ecosystems, and bridge data decentralization movements with economic empowerment.

ACKNOWLEDGMENTS

T.K.M. reports funding support from the BlockLAB at the San Diego Supercomputer Center. A.J.C. reports funding from the UC San Diego Institute for Practical Ethics and Halicioğlu Data Science Institute. K.F. reports funding support from the Emerson Collective, Lumina Foundation, and the Social Science Research Council. Authors also thank Mr. Ken Miyachi for his helpful comments and suggestions regarding this manuscript.

AUTHOR CONTRIBUTIONS

All authors wrote the manuscript and approved the final version of the manuscript.

DECLARATION OF INTERESTS

T.K.M. is the co-founder and CEO of S-3 Research, a company currently funded by the National Institutes of Health – National Institute on Drug Abuse and the U.S. Food and Drug Administration. J.Y., K.S.T., and K.F. are board members of the Native BioData Consortium.

REFERENCES

National Institutes of Health (2021). All of Us Research Program Tribal Consultation Final Report. Accessed 24 February 2022. <https://allofus.nih.gov/all-us-research-program-tribal-consultation-final-report>.

Around Him, D., Aguilar, T.A., Frederick, A., Larsen, H., Seiber, M., and Angal, J. (2019). Tribal IRBs: A Framework for Understanding Research Oversight in American Indian and Alaska Native Communities. *Am Indian Alsk Native Ment Health Res* 26, 71–95. <https://doi.org/10.5820/aian.2602.2019.71>.

Bari, L., and O'Neil, D. (2019). "Rethinking Patient Data Privacy In The Era Of Digital Health". *Health Affairs Blog* December 12. <https://doi.org/10.1377/hblog20191210.216658>.

Claw, K.G., Anderson, M.Z., Begay, R.L., Tsosie, K.S., Fox, K., and Garrison, N.A.; Summer internship for Indigenous peoples in Genomics SING

Consortium (2018). A framework for enhancing ethical genomic research with Indigenous communities. *Nat. Commun.* 9, 2957. <https://doi.org/10.1038/s41467-018-05188-3>.

Ekblaw, A., and Azaria, A. (2016). MedRec: Medical Data Management on the Blockchain. *Viral Communications*. <https://viral.media.mit.edu/pub/medrec>.

Fox, K. (2020). The Illusion of Inclusion - The "All of Us" Research Program and Indigenous Peoples' DNA. *N. Engl. J. Med.* 383, 411–413. <https://doi.org/10.1056/NEJMp1915987>.

Harding, A., Harper, B., Stone, D., O'Neill, C., Berger, P., Harris, S., and Donatuto, J. (2012). Conducting research with tribal communities: sovereignty, ethics, and data-sharing issues. *Environ. Health Perspect.* 120, 6–10. <https://doi.org/10.1289/ehp.1103904>.

Kinchin, I., Mccalman, J., Bainbridge, R., Tsey, K., and Lui, F.W. (2017). Does Indigenous health research have impact? A systematic review of reviews. *Int. J. Equity Health* 16, 52. <https://doi.org/10.1186/s12939-017-0548-4>.

Melroy-Greif, W.E., Wilhelmsen, K.C., and Ehlers, C.L. (2016). Genetic variation in FAAH is associated with cannabis use disorders in a young adult sample of Mexican Americans. *Drug Alcohol Depend.* 166, 249–253. <https://doi.org/10.1016/j.drugalcdep.2016.06.021>.

Miyachi, K., and Mackey, T.K. (2021). hOCBS: a privacy-preserving blockchain framework for healthcare data leveraging an on-chain and off-chain system design. *Inf. Process. Manag.* 58, 102535. <https://doi.org/10.1016/j.ipm.2021.102535>.

Tandon, A., Dhir, A., Islam, A.K.M., and Mäntymäki, M. (2020). Blockchain in healthcare: a systematic literature review, synthesizing framework and future research agenda. *Computers in Industry* 122, 103290. <https://doi.org/10.1016/j.compind.2020.103290>.

Tengo, M., Austin, B.J., Danielsen, F., and Fernández-Llamazares, Á. (May 2021). Creating Synergies between Citizen Science and Indigenous and Local Knowledge. *Bioscience* 71, 503–518. <https://doi.org/10.1093/biosci/biab023>.

Tsosie, K.S., Claw, K., and Garrison, N.A. (2021). Considering "Respect for Sovereignty" Beyond the Belmont Report and the Common Rule: Ethical and Legal Implications for American Indian and Alaska Native Peoples. *Am. J. Bioeth.* 21, 27–30. <https://doi.org/10.1080/15265161.2021.1968068>.

Walker, J.D., Pyper, E., Jones, C.R., Khan, S., Chong, N., Legge, D., Schull, M.J., and Henry, D. (2018). Unlocking First Nations health information through data linkage. *Int. J. Popul. Data Sci.* 3, 450. Published 2018 May 22. <https://doi.org/10.23889/ijpds.v3i1.450>.

Yellow Horse, A.J., and Huysen, K.R. (2021). Indigenous data sovereignty and COVID-19 data issues for American Indian and Alaska Native Tribes and populations. *J. Popul. Res.* 1–5. <https://doi.org/10.1007/s12546-021-09261-5>.