



IEEE Recommended Practice for Provenance of Indigenous Peoples' Data

IEEE Society of Social Implications of Technology (SSIT/SC)

Developed by the
Social Implications of Technology Standards Committee

IEEE Std 2890™-2025

IEEE Recommended Practice for Provenance of Indigenous Peoples' Data

Developed by the

Social Implications of Technology Standards Committee
of the
IEEE Society of Social Implications of Technology (SSIT/SC)

Approved 10 September 2025

IEEE SA Standards Board

Abstract: A common set of parameters for the provenance of Indigenous Peoples' data is established in this recommended practice. This recommended practice supports the appropriate disclosure of Indigenous Peoples' relationships and/or links to all data. The provenance of Indigenous Peoples' data includes information that facilitates Indigenous Peoples' governance, decision-making, participation, collaboration, and engagement in, and for, current and future uses of Indigenous Peoples' data.

Keywords: CARE Principles, Indigenous cultural and intellectual property, FAIR principles, IEEE 2890™, Indigenous Data Governance, Indigenous Data Sovereignty, Indigenous Peoples, metadata, provenance

The Institute of Electrical and Electronics Engineers, Inc.
3 Park Avenue, New York, NY 10016-5997, USA

Copyright © 2025 by The Institute of Electrical and Electronics Engineers, Inc.
All rights reserved. Published 14 November 2025. Printed in the United States of America.

IEEE is a registered trademark in the U.S. Patent & Trademark Office, owned by The Institute of Electrical and Electronics Engineers, Incorporated.

PDF: ISBN 979-8-8557-2788-3 STD28393
Print: ISBN 979-8-8557-2789-0 STDPD28393

IEEE prohibits discrimination, harassment, and bullying.

For more information, visit <https://www.ieee.org/about/corporate/governance/p9-26.html>.

No part of this publication may be reproduced in any form, in an electronic retrieval system or otherwise, without the prior written permission of the publisher.

Important Notices and Disclaimers Concerning IEEE Standards Documents

IEEE Standards documents are made available for use subject to important notices and legal disclaimers. These notices and disclaimers, or a reference to this page (<https://standards.ieee.org/ipr/disclaimers/>), appear in all IEEE standards and may be found under the heading “Important Notices and Disclaimers Concerning IEEE Standards Documents.”

Notice and Disclaimer of Liability Concerning the Use of IEEE Standards Documents

IEEE Standards documents are developed within IEEE Societies and subcommittees of IEEE Standards Association (IEEE SA) Board of Governors. IEEE develops its standards through an accredited consensus development process, which brings together volunteers representing varied viewpoints and interests to achieve the final product. IEEE standards are documents developed by volunteers with scientific, academic, and industry-based expertise in technical working groups. Volunteers involved in technical working groups are not necessarily members of IEEE or IEEE SA and participate without compensation from IEEE. While IEEE administers the process and establishes rules to promote fairness in the consensus development process, IEEE does not independently evaluate, test, or verify the accuracy of any of the information or the soundness of any judgments contained in its standards.

IEEE makes no warranties or representations concerning its standards, and expressly disclaims all warranties, express or implied, concerning all standards, including but not limited to the warranties of merchantability, fitness for a particular purpose and non-infringement. IEEE Standards documents do not guarantee safety, security, health, or environmental protection, or compliance with law, or guarantee against interference with or from other devices or networks. In addition, IEEE does not warrant or represent that the use of the material contained in its standards is free from patent infringement. IEEE Standards documents are supplied “AS IS” and “WITH ALL FAULTS.”

Use of an IEEE standard is wholly voluntary. The existence of an IEEE standard does not imply that there are no other ways to produce, test, measure, purchase, market, or provide other goods and services related to the scope of the IEEE standard. Furthermore, the viewpoint expressed at the time a standard is approved and issued is subject to change brought about through developments in the state of the art and comments received from users of the standard.

In publishing and making its standards available, IEEE is not suggesting or rendering professional or other services for, or on behalf of, any person or entity, nor is IEEE undertaking to perform any duty owed by any other person or entity to another. Any person utilizing any IEEE Standards document should rely upon their own independent judgment in the exercise of reasonable care in any given circumstances or, as appropriate, seek the advice of a competent professional in determining the appropriateness of a given IEEE standard.

IN NO EVENT SHALL IEEE BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO: THE NEED TO PROCURE SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE PUBLICATION, USE OF, OR RELIANCE UPON ANY STANDARD, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE AND REGARDLESS OF WHETHER SUCH DAMAGE WAS FORESEEABLE.

Translations

The IEEE consensus balloting process involves the review of documents in English only. In the event that an IEEE standard is translated, only the English language version published by IEEE is the approved IEEE standard.

Use by artificial intelligence systems

In no event shall material in any IEEE Standards documents be used for the purpose of creating, training, enhancing, developing, maintaining, or contributing to any artificial intelligence systems without the express, written consent of IEEE SA in advance. “Artificial intelligence” refers to any software, application, or other system that uses artificial intelligence, machine learning, or similar technologies, to analyze, train, process, or generate content. Requests for consent can be submitted using the Contact Us form.

Official statements

A statement, written or oral, that is not processed in accordance with the IEEE SA Standards Board Operations Manual is not, and shall not be considered or inferred to be, the official position of IEEE or any of its committees and shall not be considered to be, or be relied upon as, a formal position of IEEE or IEEE SA. At lectures, symposia, seminars, or educational courses, an individual presenting information on IEEE standards shall make it clear that the presenter’s views should be considered the personal views of that individual rather than the formal position of IEEE, IEEE SA, the Standards Committee, or the Working Group. Statements made by volunteers may not represent the formal position of their employer(s) or affiliation(s). News releases about IEEE standards issued by entities other than IEEE SA should be considered the view of the entity issuing the release rather than the formal position of IEEE or IEEE SA.

Comments on standards

Comments for revision of IEEE Standards documents are welcome from any interested party, regardless of membership affiliation with IEEE or IEEE SA. However, **IEEE does not provide interpretations, consulting information, or advice pertaining to IEEE Standards documents.**

Suggestions for changes in documents should be in the form of a proposed change of text, together with appropriate supporting comments. Since IEEE standards represent a consensus of concerned interests, it is important that any responses to comments and questions also receive the concurrence of a balance of interests. For this reason, IEEE and the members of its Societies and subcommittees of the IEEE SA Board of Governors are not able to provide an instant response to comments or questions, except in those cases where the matter has previously been addressed. For the same reason, IEEE does not respond to interpretation requests. Any person who would like to participate in evaluating comments or revisions to an IEEE standard is welcome to join the relevant IEEE SA working group. You can indicate interest in a working group using the Interests tab in the Manage Profile & Interests area of the [IEEE SA myProject system](#).¹ An IEEE Account is needed to access the application.

Comments on standards should be submitted using the [Contact Us](#) form.²

¹ Available at: <https://development.standards.ieee.org/myproject-web/public/view.html#landing>.

² Available at: <https://standards.ieee.org/about/contact/>.

Laws and regulations

Users of IEEE Standards documents should consult all applicable laws and regulations. Compliance with the provisions of any IEEE Standards document does not constitute compliance to any applicable regulatory requirements. Implementers of the standard are responsible for observing or referring to the applicable regulatory requirements. IEEE does not, by the publication of its standards, intend to urge action that is not in compliance with applicable laws, and these documents may not be construed as doing so.

Data privacy

Users of IEEE Standards documents should evaluate the standards for considerations of data privacy and data ownership in the context of assessing and using the standards in compliance with applicable laws and regulations.

Copyrights

IEEE draft and approved standards are copyrighted by IEEE under U.S. and international copyright laws. They are made available by IEEE and are adopted for a wide variety of both public and private uses. These include both use by reference, in laws and regulations, and use in private self-regulation, standardization, and the promotion of engineering practices and methods. By making these documents available for use and adoption by public authorities and private users, neither IEEE nor its licensors waive any rights in copyright to the documents.

Photocopies

Subject to payment of the appropriate licensing fees, IEEE will grant users a limited, non-exclusive license to photocopy portions of any individual standard for company or organizational internal use or individual, non-commercial use only. To arrange for payment of licensing fees, please contact Copyright Clearance Center, Customer Service, 222 Rosewood Drive, Danvers, MA 01923 USA; +1 978 750 8400; <https://www.copyright.com/>. Permission to photocopy portions of any individual standard for educational classroom use can also be obtained through the Copyright Clearance Center.

Updating of IEEE Standards documents

Users of IEEE Standards documents should be aware that these documents may be superseded at any time by the issuance of new editions or may be amended from time to time through the issuance of amendments, corrigenda, or errata. An official IEEE document at any point in time consists of the current edition of the document together with any amendments, corrigenda, or errata then in effect.

Every IEEE standard is subjected to review at least every 10 years. When a document is more than 10 years old and has not undergone a revision process, it is reasonable to conclude that its contents, although still of some value, do not wholly reflect the present state of the art. Users are cautioned to check to determine that they have the latest edition of any IEEE standard.

In order to determine whether a given document is the current edition and whether it has been amended through the issuance of amendments, corrigenda, or errata, visit [IEEE Xplore](#) or [contact IEEE](#).³ For more information about the IEEE SA or IEEE's standards development process, visit the IEEE SA Website.

Errata

Errata, if any, for all IEEE standards can be accessed on the [IEEE SA Website](#).⁴ Search for standard number and year of approval to access the web page of the published standard. Errata links are located under the Additional Resources Details section. Errata are also available in [IEEE Xplore](#). Users are encouraged to periodically check for errata.

Patents

IEEE standards are developed in compliance with the [IEEE SA Patent Policy](#).⁵

Attention is called to the possibility that implementation of this standard may require the use of subject matter covered by patent rights. By publication of this standard, no position is taken by the IEEE with respect to the existence or validity of any patent rights in connection therewith. If a patent holder or patent applicant has filed a statement of assurance via an Accepted Letter of Assurance, then the statement is listed on the IEEE SA Website at <https://standards.ieee.org/about/sasb/patcom/patents/>. Letters of Assurance may indicate whether the Submitter is willing or unwilling to grant licenses under patent rights without compensation or under reasonable rates, with reasonable terms and conditions that are demonstrably free of any unfair discrimination to applicants desiring to obtain such licenses.

Essential Patent Claims may exist for which a Letter of Assurance has not been received. The IEEE is not responsible for identifying Essential Patent Claims for which a license may be required, for conducting inquiries into the legal validity or scope of Patents Claims, or determining whether any licensing terms or conditions provided in connection with submission of a Letter of Assurance, if any, or in any licensing agreements are reasonable or non-discriminatory. Users of this standard are expressly advised that determination of the validity of any patent rights, and the risk of infringement of such rights, is entirely their own responsibility. Further information may be obtained from the IEEE Standards Association.

IMPORTANT NOTICE

Technologies, application of technologies, and recommended procedures in various industries evolve over time. The IEEE standards development process allows participants to review developments in industries, technologies, and practices, and to determine what, if any, updates should be made to the IEEE standard. During this evolution, the technologies and recommendations in IEEE standards may be implemented in ways not foreseen during the standard's development. IEEE standards development activities consider research and information presented to the standards development group in developing any safety recommendations. Other information about safety practices, changes in technology or technology implementation, or impact by peripheral systems also may be pertinent to safety considerations during implementation of the standard. Implementers and users of IEEE Standards documents are responsible for determining and complying with all appropriate safety, security, environmental, health, data privacy, and interference protection practices and all applicable laws and regulations.

³ Available at: <https://ieeexplore.ieee.org/browse/standards/collection/ieee>.

⁴ Available at: <https://standards.ieee.org/standard/>.

⁵ Available at: <https://standards.ieee.org/about/sasb/patcom/materials/>.

Participants

At the time this recommended practice was completed, the Indigenous Data Working Group had the following membership:

Stephanie Carroll, *Chair*
Jane Anderson, *Vice Chair*

Camille Callison, *Indigenous Caucus Chair*
Maui Hudson, *Indigenous Caucus Vice Chair*

Randall Akee
Stacy Allison-Cassin
Lauren Bennett
Kerry Blinco
Chris Chaney
Pyrrou Chung
Erica Dias
Adam Fernandez
Allison Fish
Mercury Fox

Ulia Gosart
Victoria (Vicky) Hailey
Preston Hardison
Kyla Larson
Jonathan Pringle
James Rose
Katharina Ruckstuhl
Lea Shanley
Brad Sherman

Andrew Sporle
Beckett Sterner
Gopal Tadeballi
Tran Tran
Rebecca Tsosie
Shona van Zijl de Jong
Stephanie von Gavel
Dennis Walworth
Christina Wasson
Joseph Yracheta

The following members of the individual Standards Association balloting group voted on this recommended practice. Balloters may have voted for approval, disapproval, or abstention.

Randall Akee
Jane Anderson
Matias Aranguiz Villagran
Manomit Bal
Kerry Blinco
Camille Callison
Stephanie Carroll
Erica Dias
Hassan El Shazly

A. Ferraro
Mercury Fox
Victoria Hailey
Werner Hoelzl
Maui Hudson
Piotr Karocki
Quist-Aphetsi Kester
Siri Jodha Khalsa

Weili Liu
Jonathan Mason
Monique Morrow
Bansi Patel
Louis Randazzo III
R. K. Rannow
Jhony Sembiring
Beckett Sterner
Stephanie von Gavel

When the IEEE SA Standards Board approved this recommended practice on 10 September 2025, it had the following membership:

Lei Wang, *Chair*
Jon Walter Rosdahl, *Vice Chair*
David J. Law, *Past Chair*
Alpesh Shah, *Secretary*

Edward Au
Ted Burse
Xiaofeng (Alfred) Chen
Doug Edwards
Nehad El-Sherif
J. Travis Griffith
Deborah R. Hagar

Guido R. Hiertz
Ronald W. Hotchkiss
Tyler L. Jaynes
Thomas Koshy
Howard Li
Xiaohui Liu
Kevin W. Lu
Hiroshi Mano

Daleep C. Mohla
Annette D. Reilly
Robby Robson
Daniel Sabin
F. Keith Waters
Sha Wei
Luyang (Eric) Zhan

Introduction

This introduction is not part of IEEE Std 2890™-2025, IEEE Recommended Practice for Provenance of Indigenous Peoples' Data.

This recommended practice outlines the guidelines for describing and recording the provenance of Indigenous Peoples' data. This recommended practice outlines the core parameters for providing and digitally embedding provenance information for Indigenous Peoples' data. The recommended practice establishes common descriptors and controlled vocabulary for provenance, including recommendations for metadata fields that can be used across industry sectors, including machine learning (ML) and artificial intelligence (AI) contexts, biodiversity and genomic science innovation, and other associated socio-economic and/or administrative databases from various academic, governmental, and non-profit sectors. This recommended practice supports proper and appropriate disclosure of originating data information, as well as the long-term identification of Indigenous Peoples' data for future use, connecting data to people and place, and, when appropriate, supporting future benefit-sharing options.

Background

Following the 2007 United Nations Declaration on the Rights of Indigenous Peoples (UNDRIP), this recommended practice acknowledges the sovereign right of Indigenous Peoples to self-identify and to determine governance and decision-making processes for membership and/or citizenship. This document also recognizes the varied contexts, histories, and politics around the world that may result in a collective group with ancestral ties to lands to not identify as Indigenous, particularly communities outside settler colonial states. In some parts of Asia and Africa, state governments use the term “ethnic groups” or “ethnic minorities,” while these groups themselves identify as Indigenous. Within any one “Indigenous People,” there may be many subgroups. This terminology does not address the vast diversity of how Indigenous Peoples organize themselves.

Historical collections of Indigenous data currently held in archives, libraries, museums, by governments and corporations, and in research contexts teach important lessons about failures in proper documentation, classification, and cataloguing practices. During the transition of these collections into digital formats, the lack of proper practices has resulted in significant errors in metadata for Indigenous Peoples' data. As a consequence, Indigenous data is difficult to find, access, and reuse, and is often not interoperable. This issue limits Indigenous Peoples' own use of the data as well as their decision-making and governance over how, when, and why these data are to be used. Thus, Indigenous Peoples' benefits from and contributions to the value of data for society are dependent on addressing impoverished or absent data and metadata.

Additionally, there has been significant misuse of Indigenous Peoples' data over time and space. These uses have resulted in social, cultural, environmental, and economic harm to Indigenous Peoples. There have also been predatory practices that have targeted Indigenous cultural and intellectual property. The misappropriation, commodification, and publication of some of these data and related content have occurred within and across nation-states. The absence of clear rules and standards has been an important obstacle in preventing these types of abuses.

Indigenous cultural and intellectual property means the cultural heritage of Indigenous Peoples, which comprises all objects, sites, and knowledge, the nature and use of which have been transmitted or continue to be transmitted from generation to generation, and which is regarded as pertaining to a particular Indigenous group or its territory. This cultural heritage is a living concept that can be adaptive and includes objects, knowledge, literary and artistic works, which may also be created in the future based on that cultural heritage, and includes the following:

- Literary, performing, and artistic works (including songs, music, dances, stories, ceremonies, symbols, languages, and designs)
- Scientific, agricultural, technical, and ecological knowledge (including about cultigens, medicines, and phenotypes of flora and fauna)

- All items of movable and immovable cultural property (including sacred and historically significant sites and burial grounds) and ancestral remains
- Documentation of Indigenous Peoples' heritage in archives, film, photographs, videotape, or audiotape, and in all forms of media

Indigenous Peoples' data is a form of Indigenous cultural and intellectual property where that data relates to any of the previous list items.

Across the world, there are significant decentralized activities where Indigenous Peoples, their nations, leaders, advocates, and communities are actively exercising and using various Indigenous data for governance, teaching, archiving, and restorative work. There is also an increasing number of Indigenous data ethics and governance regulations, protocols, principles, and preferences across scales. Examples include the high-level CARE (Collective benefit, Authority to control, Responsibility, Ethics) Principles for Indigenous Data Governance; the regional Asia Framework on Indigenous Knowledge and Data Sovereignty; the Protocols for Native American Archival Materials in the United States; and Indigenous Peoples' own laws, regulations, policies, and preferences that require enriched metadata. These various activities occur at all levels of society and necessitate an international standard for Indigenous Peoples' data provenance.

One way to address these problems is to develop new standards for the appropriate inclusion of Indigenous provenance in data and metadata fields across multiple research, institutional, and industry sectors. An additional motivation for the development of the recommended practice is to support the clear disclosure of the origin of Indigenous interests in national and international patent office searches, particularly in the biodiversity and genomic innovation sectors. This practice fosters Indigenous leadership and participation in decisions about the use of Indigenous data, as well as the opportunity for direct benefits to flow back to Indigenous Peoples and communities. Research sectors also benefit from having access to more accurate and comprehensive information about data sources. The market tends to respond positively to greater provenance and transparency around data, as these qualities increase trust in the data itself.

Within data-generated industries, such as artificial intelligence and machine learning, as well as biodiversity research and innovations in genomic science, an Indigenous data standard will enhance the integrity of that data, as well as the likelihood of engaged collaborations with data producers in ways that were previously unthinkable. This can help build equitable partnerships in the protection, use, and application of the data, including commercial partnerships and benefit sharing. It can also support the development of consumer and public trust across a wide spectrum of data industries. Setting standards begins to address the combined problems of the proliferation of unreliable data and information on Indigenous Peoples; the lack of Indigenous participation, control, and governance over the circulation of Indigenous data and information; and the misappropriation and misuse of Indigenous data and information.

Acknowledgments

Sincere thanks are extended to all the Indigenous community representatives and Indigenous scholars who were able to join the special webinars and contributed to the draft. Gratitude is also expressed to those who made contributions in the working groups or at other stages in the five-year process.

The following individuals, representing distinct subgroups described in [Annex B](#), are recognized for their valuable contributions:

- Conceptual subgroup: Maui Hudson and Adam Fernandez
- Scope and Purpose subgroup: Randy Akee and Tran Tran
- Standard Landscapes subgroup: Stacy Allison-Cassin and Stephanie von Gavel
- Drafting subgroup: Christopher Chaney, Brad Sherman, Stephanie Carroll, Jane Anderson, and Randy Akee

Contents

1. Overview	11
1.1 Scope	11
1.2 Purpose	11
1.3 Word usage	11
2. Normative references	12
3. Definitions	12
4. Approach and recommended practice for provenance of Indigenous Peoples' data	13
4.1 General	13
4.2 Indigenous relationships and/or links to data	13
4.3 Indigenous Peoples' data	13
4.4 Data actors	14
5. Disclosure of Indigenous Peoples' relationships and/or links to data	14
6. Inclusions for provenance of Indigenous Peoples' data	15
6.1 General	15
6.2 Identification information	15
6.3 Data details	15
6.4 Data fields	16
7. Using this recommended practice for the provenance of Indigenous Peoples' data	16
7.1 General	16
7.2 Non-human data actors	16
7.3 Application of this recommended practice for the provenance of Indigenous Peoples' data	17
8. When pre-existing data does not comply with this recommended practice for the provenance of Indigenous Peoples' data	18
9. Multiple interests in data	18
10. Applicable timeframe for this recommended practice	19
11. Where to disclose provenance information for Indigenous Peoples' data	19
12. Format to disclose provenance information for Indigenous Peoples' data	19
13. Timing of disclosure of provenance information for Indigenous Peoples' data	20
13.1 General	20
13.2 When does the obligation end?	20
14. Limits on disclosure of provenance information	20
Annex A (informative) Bibliography	21
Annex B (informative) Subgroups of the Working Group	22
Annex C (informative) Indigenous Peoples' Caucus	23

IEEE Recommended Practice for Provenance of Indigenous Peoples' Data

1. Overview

1.1 Scope

This recommended practice details the process by which the provenance of Indigenous Peoples' data should be described and recorded. This recommended practice outlines the core parameters for providing and digitally embedding the provenance of Indigenous Peoples' data. This recommended practice supports appropriate disclosure of Indigenous Peoples' relationships and/or links to data, as well as the long-term identification of Indigenous Peoples' data for future use, and connects data to people and place, and when appropriate, facilitates future benefit sharing.

1.2 Purpose

The purpose of this recommended practice is to describe what provenance of Indigenous Peoples' data includes; how the recommended practice should be operationalized; when, where, and how to disclose the provenance information and in what format; as well as what, if any, potential limitations exist for recognizing the provenance of Indigenous Peoples' data. This recommended practice aims to facilitate the interoperability of provenance fields across various information platforms and databases and to support the disclosure and use of Indigenous Peoples' data.

1.3 Word usage

The word *shall* indicates mandatory requirements strictly to be followed in order to conform to the standard and from which no deviation is permitted (*shall* equals *is required to*).^{6,7}

The word *should* indicates that among several possibilities one is recommended as particularly suitable, without mentioning or excluding others; or that a certain course of action is preferred but not necessarily required (*should* equals *is recommended that*).

The word *may* is used to indicate a course of action permissible within the limits of the standard (*may* equals *is permitted to*).

⁶ The use of the word *must* is deprecated and cannot be used when stating mandatory requirements; *must* is used only to describe unavoidable situations.

⁷ The use of *will* is deprecated and cannot be used when stating mandatory requirements; *will* is only used in statements of fact.

The word *can* is used for statements of possibility and capability, whether material, physical, or causal (*can* equals *is able to*).

2. Normative references

This document does not contain any normative references.

3. Definitions

For the purposes of this document, the following terms and definitions apply. *The IEEE Standards Dictionary Online* should be consulted for terms not defined in this clause.⁸

audit trail: A detailed chronological record of all activities related to data or reports, including creation, modification, relocation, and deletion.

data: A reinterpretable formal representation of information for communication, interpretation, or processing. Possible: 1) Any representations, such as characters or analog quantities, to which meaning might be assigned. 2) Representation of facts, concepts, or instructions in a formalized manner suitable for communication, interpretation, or processing by human beings or by automatic means (IEEE Std 2413-2019).

NOTE—Data can be processed manually or automatically.⁹

data actor: A person, device, application, system, or organization that communicates, generates, uses, or stores information or data.

NOTE—Based on the definition of actor in IEEE Std 9274.1.1-2023 [B1].¹⁰

derogatory treatment: Use of Indigenous Peoples' data in a way that undermines the originally intended meaning of those symbols, iconography, and/or narratives.

Indigenous cultural and intellectual property rights: Moral and legal rights and interests that Indigenous Peoples assert in order to protect Indigenous Peoples' data, including from **derogatory treatment**, **misappropriation**, **misattribution**, and **misuse**.

NOTE—Indigenous Peoples assert the moral right to integrity and to be recognized, attributed, acknowledged, named, and perpetually associated with Indigenous Peoples' data.

Indigenous Peoples: A collective group with ancestral ties to lands.

NOTE—This terminology does not address the vast diversity of how Indigenous Peoples organize or identify themselves.

Indigenous Peoples' data: Data, information, and knowledges recorded on any medium and in any format generated by Indigenous Peoples as well as by governments, the private sector, and other institutions or individuals on and about Indigenous Peoples, their governments, or more-than-human relations.

⁸IEEE Standards Dictionary Online is available at: <http://dictionary.ieee.org>. An IEEE account is required for access to the dictionary, and one can be created at no charge on the *dictionary sign-in page*.

⁹Notes in text, tables, and figures of a standard are given for information only and do not contain requirements needed to implement this standard.

¹⁰Numbers in brackets correspond to the sources in Annex A.

Indigenous sovereignty: Inherent rights derived from spiritual ways, culture, language, social and legal systems, political structures, and inherent relationships with lands, waters, and all upon them.

Indigenous data sovereignty: The right of Indigenous Peoples to govern the collection, ownership, and application of Indigenous Peoples' data.

misappropriation: The use of Indigenous Peoples' data for purposes without permission that are unintended by the original creator and can cause cultural, emotional, economic, and psychological harm to Indigenous Peoples.

misattribution: Missing or incorrect attribution and association of Indigenous Peoples' data.

misuse: Used improperly, not as originally intended, without consent, or in a harmful way.

NOTE—Examples of misuse include derogatory treatment, misappropriation, and misattribution.

4. Approach and recommended practice for provenance of Indigenous Peoples' data

4.1 General

This recommended practice for provenance of Indigenous Peoples' data includes information that facilitates Indigenous Peoples' governance, decision-making, participation, collaboration, and engagement in and for current and future uses of Indigenous Peoples' data.

4.2 Indigenous relationships and/or links to data

Indigenous Peoples' relationships and/or links to data arise whenever data originates from Indigenous Peoples (individually and/or collectively), their governments, their more-than-human relations, and are derived from, based upon, or generated from any of the following:

- Tangible object, material, or resource; or,
- Intangible knowledge, information, language, or culture.

Indigenous Peoples' relationships and/or links to data are supported by various articles in the United Nations Declaration on the Rights of Indigenous Peoples and have been further articulated by the Global Indigenous Data Alliance as Indigenous Peoples' Rights in Data. Data actors should comply with this recommended practice even where there are no legislated requirements or disclosed interests.

4.3 Indigenous Peoples' data

Indigenous Peoples' data refers to data, information, and knowledges recorded on any medium and in any format generated by Indigenous Peoples, as well as by governments, the private sector, and other institutions or individuals, on and about Indigenous Peoples, their governments, or their more-than-human relations. Indigenous Peoples' data is comprised of, but not limited to, information, specimens, and knowledges about more than humans with which they have relations; information about Indigenous individuals; and information and knowledges about Indigenous Peoples as collectives.

Indigenous Peoples' data includes, but is not limited to, the following:

- Writings (e.g., a book, poetry)
- Music (e.g., a song)
- Performances (e.g., a dance, a spoken reading)
- Ceremonies
- Artistic work (e.g., a painting, sculpture, design, symbols/motifs)
- Languages
- Oral traditions and cultural narratives
- Documentation of Indigenous Peoples' heritage in all forms of media (e.g., reports, photography, films, sound recordings)
- Specimens (e.g., plant, soil, living or non-living organisms)
- Individual-level socio-economic data and all health-related data and specimens
- Administrative data from various governmental agencies, nations, and organizations
- Environmental data

4.4 Data actors

Data actors are a person, device, application, system, or organization that communicates, generates, uses, or stores information or data, including but not limited to the following:

- Researchers and research institutions
- Repositories and collections
- Databanks
- Aggregators
- Publishers
- Machines, algorithms, and other sensor, observation, and artificial intelligence technology
- Data scrapers

The definition of *data actor* applies equally to public, private, governmental, administrative, commercial, and non-commercial actors.

5. Disclosure of Indigenous Peoples' relationships and/or links to data

Data should be disclosed whenever a data actor is made aware or becomes aware that Indigenous Peoples are in relationships and/or are linked to the data that they are using. Indigenous Peoples should be consulted and their approval obtained before the tangible object, material, or resource, or the *intangible* knowledges, information, language, or culture is first collected or accessed. This should include discussions with Indigenous Peoples about what is being accessed or collected, how it is to be described, whether there are any restrictions on subsequent use (including on derived data), and the relevant Indigenous data provenance information to be disclosed.

6. Inclusions for provenance of Indigenous Peoples' data

6.1 General

The provenance of Indigenous Peoples' data should include identification information that facilitates Indigenous Peoples' governance, decision-making, participation, collaboration, and engagement in and for current and future uses of Indigenous Peoples' data.

6.2 Identification information

Provenance of Indigenous Peoples' data should include identification information about the following:

- a) The tangible object, material, or resource, or the intangible knowledges, information, language, or culture to which the provenance pertains.
- b) The Indigenous Peoples to whom the tangible object, material, or resource, or the intangible knowledges or information was generated, produced, or collected.
- c) The location where the tangible object, material, or resource, or the intangible knowledges, information, languages, or culture was sourced—either the original event of data collection or the particular copy or version being used.
- d) The date when the tangible object, material, or resource, or the intangible knowledges, information, languages, or culture was sourced—either the original event of data collection or the particular copy or version being used.
- e) The origin point of already generated data and details about its movement and location of storage over time.
- f) Any other relevant information that the Indigenous Peoples with relationships and/or links to the data want to include beyond what is listed in the five previous items as part of the provenance information for Indigenous Peoples' data.

6.3 Data details

The provenance of Indigenous Peoples' data should:

- a) Use scientific names and any names that Indigenous Peoples use.
- b) Include the following data:
 - 1) Whether permission was granted to collect the tangible object, material, or resource, or the intangible knowledges or information.
 - 2) Who granted the permission to collect the tangible object, material, or resource, or the intangible knowledges or information.
 - 3) When the tangible object, material, or resource, or the intangible knowledges or information was collected.

- 4) Messages, operations, actors, preferences, and context that led to that piece of data.
- c) Endure and persist in parallel with the context of storing data and accessibility of the data after the collection process has ended.
- d) Require a verifiable audit trail associated with the data. This may include hashes, digital signatures, or other standard practices for assuring the confidentiality, integrity, and availability of data in cybersecurity and information assurance.

6.4 Data fields

Provenance of Indigenous Peoples' data should be provided using the following types of data fields:

- Rights
- Permissions
- Author-rights
- Geo-location
- Contributors
- Citations or references
- Identifiers or links to taxonomies or metadata classifications
- Version history documentation or comments

7. Using this recommended practice for the provenance of Indigenous Peoples' data

7.1 General

This recommended practice should be used by all data actors who generate, engage with, or otherwise have decision-making power over data that discloses Indigenous Peoples' relationships and/or links.

7.2 Non-human data actors

There are situations when a machine actor, aggregator, data scraper, or other non-human data actor interacts with data that may have relationships and/or links to Indigenous Peoples. In these situations, all reasonable efforts should be made by parties who design, own, control, use, or benefit from non-human data actors to help ensure that the non-human data actor conforms with the FAIR principles and complies with the scope and operation of the Recommended Practice for Provenance of Indigenous Peoples' Data.

Reasonable efforts entail fair, proper, or appropriate actions under the circumstances and are subject to assessment by Indigenous Peoples, colleagues, peers, and the extended community of practice.

7.3 Application of this recommended practice for the provenance of Indigenous Peoples' data

7.3.1 When a data actor generates data

Whenever a data actor generates data in which Indigenous Peoples' relationships and/or links have been disclosed, the data actor should verify that the data includes provenance information for Indigenous Peoples' data.

7.3.2 When a data actor transfers data

Whenever a data actor transfers data in which Indigenous Peoples' relationships and/or links have been disclosed, the data actor should verify that:

- a) The provenance information for Indigenous Peoples' data is recorded as metadata and accompanies the data.
- b) The obligation to maintain provenance information for Indigenous Peoples' data is transferred to the data actor receiving the data.

7.3.3 When a data actor receives, collects, or stores data

Whenever a data actor receives, collects, or stores data in which Indigenous Peoples' relationships and/or links have been disclosed, the data actor should help to ensure that:

- a) Indigenous Peoples' data include provenance information.
- b) The provenance information for Indigenous Peoples' data is maintained in the system/context of storage, hosting, and use.
- c) Whenever the data actor is an institution, databank, data repository, or similar entity, the data actor should help to ensure that the infrastructure and metadata fields conforming to the FAIR Principles exist to allow provenance information for Indigenous Peoples' data to be included.

7.3.4 When a data actor uses, reuses, transforms, or processes data

Whenever a data actor uses, reuses, transforms, or processes data in which Indigenous Peoples' relationships and/or links have been disclosed, they should verify that appropriate acknowledgement and attribution are provided. The provenance information for Indigenous Peoples' data should also appear in metadata records where relevant.

7.3.5 When a data actor publishes, discloses, or otherwise makes available research or analysis that is based on data

Whenever a data actor publishes, discloses, or otherwise makes available to the public research or analysis that is based on, built on, or draws from data in which Indigenous Peoples' relationships and/or links have been disclosed, they should help to ensure that it includes relevant provenance information for Indigenous Peoples' data.

7.3.6 When a data actor destroys or disposes of data in which there is disclosed Indigenous Peoples' relationships and/or links

Whenever a data actor destroys or disposes of data in which Indigenous Peoples' relationships and/or links have been disclosed, they should gain permission to dispose from the relevant communities or dispose at the request of the relevant communities, and dispose of data in a manner deemed appropriate by those communities and/or provide communities with certificates of destruction.

8. When pre-existing data does not comply with this recommended practice for the provenance of Indigenous Peoples' data

When pre-existing data does not comply with this recommended practice for the provenance of Indigenous Peoples' data, data actors should ensure the following actions are taken:

- a) Where a data actor hosts pre-existing data in which there are disclosed Indigenous Peoples' relationships and/or links that do not include provenance information, the data actor should undertake due diligence to locate, identify, and include provenance information with the data, and keep a record.
- b) Where a data actor has been notified or becomes aware of Indigenous Peoples' relationships and/or links to data they are housing, storing, processing, or using, the data actor should disclose those relationships and/or links by notifying other data actors. They should also make reasonable efforts to identify and include relevant provenance information. Where such information is found, the data actor should include it with the data. As part of this process, the data actor should make reasonable efforts to consult with the relevant Indigenous Peoples regarding the appropriate provenance information and any limitations on its use.
- c) Where, after reasonable efforts to find provenance information for Indigenous Peoples' data cannot be found, the data actor can continue to use the data so long as they:
 - 1) Acknowledge disclosed Indigenous Peoples' relationships and/or links to data in a way that is visible, accessible, and fixed.
 - 2) Acknowledge that the data was collected and stored without the relevant provenance information.
 - 3) Acknowledge that reasonable efforts have been made to find the provenance information and those efforts have been unsuccessful.
 - 4) Are willing to modify the data to comply with the obligations to include provenance information if it becomes available.

9. Multiple interests in data

In some situations, tangible objects, materials, and resources, as well as intangible knowledges or information originating from Indigenous Peoples (individually and/or collectively), their governments, and/or their non-human relations are shared between different Indigenous Peoples and communities. In these situations, the data actor should:

- Acknowledge that there are shared knowledges and information.

- Include details, if available, about the different Indigenous Peoples that have disclosed relationships and/or links to the data as part of the disclosure of provenance information for Indigenous Peoples' data.

Where relevant provenance information becomes available after the data has been generated, the data actor should undertake due diligence to modify the record to include the new information.

10. Applicable timeframe for this recommended practice

When objects, materials, resources, knowledges, or information in which there are disclosed Indigenous Peoples' relationships and/or links that are modified, adapted, used in part, transformed, or combined with other objects, materials, resources, knowledges, or information, the obligations to maintain provenance information will continue to apply if the unique contribution of that data is recognizable.

The decision as to whether there is a recognizable Indigenous relationship when objects, materials, resources, knowledges, or information in which there is a disclosed Indigenous relationship are modified, adapted, used in part, or combined with other objects should be decided upon by the relevant Indigenous Peoples.

11. Where to disclose provenance information for Indigenous Peoples' data

Provenance information for Indigenous Peoples' data should be included wherever the information is used, stored, or disclosed including the following:

- Databases/banks
- Metadata records (where specific provenance fields exist, these should be used)
- Models and algorithms
- Datasets (both structured and unstructured)
- Catalogs
- Finding aids
- Survey forms used during research
- Field reports
- Publications
- Presentations
- Data transfer agreements

This provision applies equally to public, private, governmental, administrative, commercial, and non-commercial actors.

12. Format to disclose provenance information for Indigenous Peoples' data

Provenance information for Indigenous Peoples' data should:

- Not be subject to any legal, technical, or financial restrictions that limit use or access to Indigenous Peoples' data (other than limitations set by Indigenous Peoples themselves).
- Be disclosed in a manner that helps to ensure the information is visible, accessible, and stable.

Files containing provenance information for Indigenous Peoples' data should, where relevant:

- Be formatted using an open, non-proprietary file format where possible, including “.txt” or “.csv”.
- Be machine-readable and use structured fields where possible.
- Re-use existing data formats and metadata standards where possible.
- Include identifiers (e.g., DOIs, museum accession numbers, ISBNs, version numbers) where these exist unless there is an explicit reason or reasons not to do so.
- Be versioned and given their own unique identifier unless these are addressed as part of a larger file or bundle of files.
- Be attached to each item or the catalog entry for a collection when all items are of the same source.

13. Timing of disclosure of provenance information for Indigenous Peoples' data

13.1 General

Provenance information for Indigenous Peoples' data should be disclosed either before or at the time when the data are first disclosed. Where this is not possible, provenance information should be disclosed as soon as possible thereafter.

13.2 When does the obligation end?

If a certain component of Indigenous Peoples' data is used in collaboration, conjunction, or as a foundation for future data or other outputs, then this obligation to disclose the provenance information must continue on into those novel creations. The purpose is to establish a clear, direct chain of provenance information for all data related to and about Indigenous Peoples, as described earlier in [Clause 7](#).

14. Limits on disclosure of provenance information

The disclosure of provenance information for Indigenous Peoples' data should take into account and respect Indigenous Peoples' relationships and community decision-making determinations, including privacy, confidentiality, and any other reason given by Indigenous Peoples. In some cases, this means that Indigenous Peoples themselves may limit the disclosure of provenance information.

Annex A

(informative)

Bibliography

Bibliographical references are resources that provide additional or helpful material but do not need to be understood or used to implement this standard. Reference to these resources is made for informational use only.

[B1] IEEE Std 9274.1.1™-2023, IEEE Standard for Learning Technology—JavaScript Object Notation (JSON) Data Model Format and Representational State Transfer (RESTful) Web Service for Learner Experience Data Tracking and Access. (<https://standards.ieee.org/ieee/9274.1.1/7321/>)

Annex B

(informative)

Subgroups of the Working Group

B.1 Conceptual Issues subgroup

The Conceptual Issues subgroup explored the intersection between Indigenous concepts and existing understandings of provenance and related terms. As an Indigenous-focused sounding board for general issues, the subgroup covered issues as diverse as: data as relationships, data as responsibility, data as property, data as collectives, data and attribution, and data in time and space.

B.2 Scope and Purpose subgroup

The Scope and Purpose subgroup enumerated the motivation, background, and need for Indigenous data provenance standards and identified the contexts and environments in which the provenance standard should be used. This subgroup also detailed the types of data (and their uses) that are covered by the provenance standard. Finally, the subgroup identified the range of agencies and organizations that should employ these standards as well as the reporting requirements, including the need for technological requirements and harmonization across systems and time.

B.3 Standards Landscape subgroup

The Standards Landscape subgroup conducted a meta-analysis of the existing and applicable standards for review by the other working groups related to Indigenous Data Sovereignty and provenance topics.

B.4 Drafting subgroup

The Drafting subgroup set up the initial document and created the main components for the IEEE Indigenous Data Sovereignty provenance standards.

Annex C

(informative)

Indigenous Peoples' Caucus

The Indigenous Peoples' Caucus, comprised of self-identifying Indigenous people, met independently of the working group and the subgroups described previously. The members of this caucus were also participants and leaders in the various subgroups. This particular caucus discussed any potential items or concerns that would arise in the development of the IEEE Standard. This was intended as an additional forum for discussions and consideration of important items affecting the provenance of Indigenous Peoples' data. The Indigenous Peoples' Caucus, with the working group chair and vice chair, led the consultation on the recommended practice with Indigenous Peoples and individuals in 2023.

RAISING THE WORLD'S STANDARDS

Connect with us on:



Facebook: facebook.com/ieeesa



LinkedIn: linkedin.com/groups/1791118



Beyond Standards blog: beyondstandards.ieee.org



YouTube: youtube.com/ieeesa

standards.ieee.org

Phone: +1 732 981 0060