

# Advancing Data Equity: An Action-Oriented Framework

WHITE PAPER  
SEPTEMBER 2024



# Contents

|  |    |
|--|----|
| Preface  | 3  |
| Executive summary  | 4  |
| 1 Defining data equity                                   | 5  |
| 2 Implementing data equity                               | 7  |
| 3 Case studies   | 16 |
| Climate data collection and monitoring                   | 17 |
| Women's access to financial services in emerging markets | 18 |
| Racial bias in healthcare                                | 19 |
| Improving how the City of Boston collects gender data    | 20 |
| Mapping Indigenous territories                           | 21 |
| National Statistical Offices as data stewards            | 22 |
| 4 Recommendations for stakeholders                       | 23 |
| Conclusion   | 25 |
| Contributors   | 26 |
| Endnotes   | 28 |

## Disclaimer

This document is published by the World Economic Forum as a contribution to a project, insight area or interaction. The findings, interpretations and conclusions expressed herein are a result of a collaborative process facilitated and endorsed by the World Economic Forum but whose results do not necessarily represent the views of the World Economic Forum, nor the entirety of its Members, Partners or other stakeholders.

© 2024 World Economic Forum. World Economic Forum reports may be republished in accordance with the [Creative Commons AttributionNonCommercial-NoDerivatives 4.0 International Public License](#), and in accordance with our [Terms of Use](#).

## This report is interactive



Look out for this icon for pages that can be interacted with



To ensure interactive capability, please download and open this PDF with Adobe Acrobat.

# Preface



**JoAnn Stonier**  
Mastercard Fellow,  
Data and AI, Mastercard



**Lauren Woodman**  
Chief Executive Officer,  
DataKind

In 2023, as generative artificial intelligence (genAI) and other technologies expanded their adoption and impact on society, the World Economic Forum established a Global Future Council focusing on the issue of data equity.<sup>1</sup> Through research and discussions with experts in technology, data, business and social science, it became clear that a foundational definition and approach needed to be created to allow organizations of all types to build more equitable systems, processes, practices and outcomes. Our initial thoughts on this topic were published in our first white paper titled “Data Equity: Foundational Concepts for Generative AI”.<sup>2</sup>

As our work evolved, it became clear that our data-driven world was not created in a manner that drives equitable outcomes, simply because it was not designed with equity in mind. It was created with all of our societal varieties, historical inequities, biases and differences. While we want these differences to be reflected in our technological solutions, we do not want those differences perpetuated, amplified or extended in our technology solutions. We want technology to create a better and more inclusive future, one where we solve problems, not repeat past ones.

Our research and consultations revealed that data equity impacts diverse sectors, industries and

regions. This complexity necessitates a flexible approach. In response, we developed a framework for responsible data practices that adapts to specific contexts while ensuring consistency and compliance with global regulations.

The present white paper provides the global community with a baseline definition and a data equity framework for inquiry to be used as a guide to help spur conversations and self-assessment inside organizations as they seek to use AI more broadly. This report builds on the four types of equity proposed in our initial briefing paper (representation, feature, access and outcome equity)<sup>3</sup> and proposes 10 characteristics that need to be considered by organizations as they build out systems, products and solutions via a framework for action.

It is our hope that as organizations utilize our data equity definition and framework, the issues and considerations required for equitable outcomes will become clear. It is our recommendation that all organizations, no matter their nature (commercial, civil society, academic or governmental), recognize that we must increase our understanding and improve our design methodology in order to design a future that ensures outcomes for a balanced and equity-driven world.

# Executive summary

Data equity requires collective action throughout the data life cycle to ensure fair, just and beneficial outcomes for all.

Data equity is a shared responsibility that requires collective action to create data practices and systems that promote fair and just outcomes for all. Continuously considering the human impact of data is of critical importance given the ever-expanding role of data-driven systems in today's increasingly digital societies. By considering data equity throughout the data life cycle, data practices can be improved to promote fair, just and beneficial outcomes for all individuals, groups and communities.

The Global Future Council (GFC) on Data Equity, a multistakeholder group of experts, has come together during 2023-2024 to define and create a "framework of inquiry" for data equity. This data equity framework is designed to prompt reflection, focus research and guide corrective action. This unique framework offers a culturally-grounded perspective on data management and governance. It is based on the Te Mana o te Raraunga Model, a Māori data sovereignty model that describes the

internal logic that traditional knowledge-keepers use when deciding to share knowledge with others. Additionally, the framework is aligned with existing data governance guidelines and principles, including FAIR, CARE, TRUST, to demonstrate how data equity complements existing modalities and enriches the broader discussion of the appropriate use of data in modern life.

The framework consists of 10 characteristics and related key issues, grouped into three main categories: data, purpose and people. As part of the framework, a series of questions have been developed to evaluate data and initial actions suggested to guide stakeholders in implementing data equity in their organizations. Though this framework is rooted in Indigenous data sovereignty, it provides guidance and encourage reflection for advancing data equity across sectors, communities and geographies. Six case studies demonstrate the use of the framework through real-world examples, and serve as inspiration for putting this tool into practice in other contexts.

1

# Defining data equity

A shared definition of data equity is essential to advance collaboration and coordinated action to put this concept into practice.

Advancing data equity is essential. We live in an era where automated decision-making systems based on algorithms and data are increasingly common, with profound implications for individuals, communities and society. Those designing and using such systems must carefully consider the potential social impact, with all-round equity as a core concern.

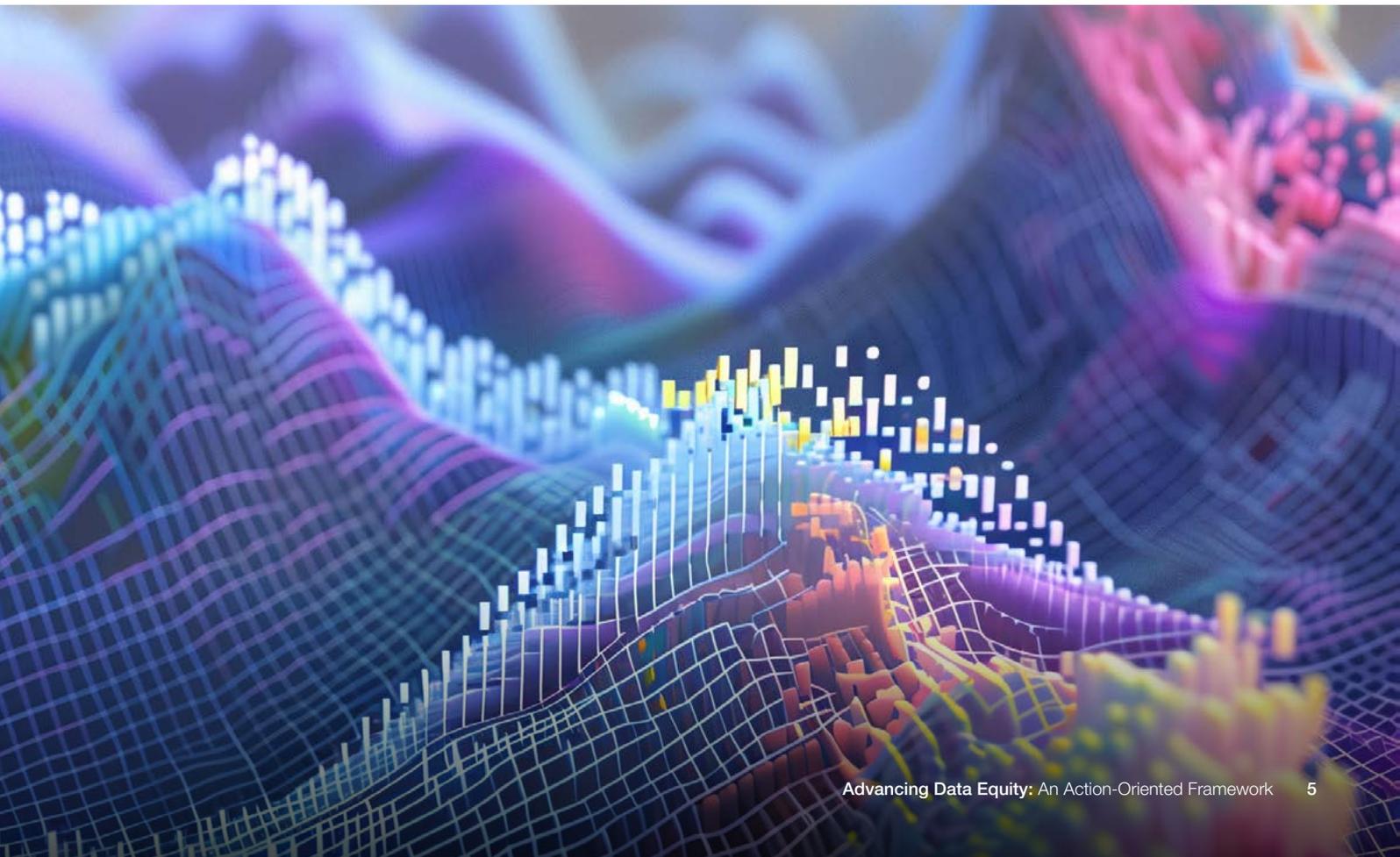
Despite its burgeoning significance, the concept of “data equity” lacks a clear, widely accepted definition in policy circles and academic literature.<sup>4</sup> Perhaps the most widely-known definition defines data equity as the social concept of fairness applied to computer science and machine learning, and identifies various aspects of data equity, including representation, feature, access and outcome equity.<sup>5</sup>

This ambiguity does not only impede progress but also risks exacerbating the very inequities that stakeholders aim to address. Without a shared understanding, stakeholders are left to interpret and implement data equity measures based on their own, potentially conflicting, perspectives, and without a clear benchmark against which to

measure their efforts. Moreover, as technological advancements accelerate and data becomes increasingly critical, new challenges to data equity continue to emerge. And on a global scale, the absence of a common understanding hampers international collaboration on this crucial issue.

Recognizing this critical gap, the GFC on Data Equity has crafted a comprehensive definition to foster alignment and drive meaningful progress:

**Data equity** is the shared responsibility for fair data practices that respect and promote human rights, opportunity and dignity. Data equity is a fundamental responsibility that requires strategic, participative, inclusive and proactive collective and coordinated action to create a world where data-based systems promote fair, just and beneficial outcomes for all individuals, groups and communities. It recognizes that data practices – including collection, curation, processing, retention, analysis, stewardship and responsible application of resulting insights – significantly impact human rights and the resulting access to social, economic, natural and cultural resources and opportunities.<sup>6</sup>



Data equity seeks to address historical, current and potential imbalances in datasets that are used in a variety of domains in data-driven decisions and algorithmic and AI systems. In addition, data equity is concerned with access to datasets as well as how, and by whom, they are used in societally impactful decision-making and systems. Participative and collective responsibility and decision-making, especially by individual and collective data subjects, is a central tenet. Therefore, data equity serves as the foundation of fairness and justness in the development and application of a host of technologies and for earning trust for digital systems – as described in the World Economic Forum’s Digital Trust Framework.<sup>7</sup>

Data equity can be achieved by appropriate design of data collection, uses, practices and governance in order to promote just and fair outcomes for people and communities directly or indirectly impacted by these systems. In this regard, the focus of data studies must adapt to include not just what “data is”, but also what “data does”. The proposed definition, while covering the whole “data life cycle,” particularly centres on the impact side of data governance and practices.

Data equity considerations permeate the whole data life cycle, for example: how data is collected and constructed (input data equity); made available (data access equity); made representative and relevant for the context and purpose it is being used (data representation equity); processed and interpreted (process or algorithmic data equity); used to generate and inform outcomes (outcome

data equity); and how its value is being distributed and shared with individuals and communities that have contributed to it (data value equity).<sup>8</sup>

Thus, it is crucial to consider data equity from the earliest stages of the data life cycle, as quality and equity issues might not be easily remedied later. Moreover, data collectors should also consider the possible subsequent (re)use of their data by other actors in potentially harmful or exploitative ways.

Data equity can be advanced through **corrective** as well as **proactive** actions in the different stages of the data life cycle:

- **Corrective actions** include addressing historical (and current or potential) biases in datasets, such as biased depictions or under-representation of marginalized groups, as well as giving individuals and communities the ability to control their own data (through opt-in or opt-out mechanisms) in order to ensure their individual and collective agency, autonomy and right to privacy.
- **Proactive actions** include engaging individuals represented in the data to help define it; employing collection methods that enable identification, representation and participation of diverse groups; promoting open and transparent data sharing; developing inclusive, participatory systems that utilize the data, ensuring those affected have a voice; verifying that these systems produce fair and equitable outcomes; and guaranteeing that data contributors benefit from the value generated by its use.

# Implementing data equity

The data equity framework is designed to encourage reflection, guide research and prompt corrective actions.

“Essentially, this framework should be regarded as a “framework for inquiry”, i.e. a guide to help spur conversations and evaluation inside organizations and communities as they consider using data, whether in AI-enabled systems or elsewhere.

At its core, the pursuit of equity is about uplifting people and ensuring just and fair treatment for all. While the concept of data equity is relatively new, its application in the context of genAI intersects with long-existing issues relevant to data governance, trustworthiness, privacy and responsible data use. Addressing equity in these and other data-related issues involves technical considerations, but their explicit human and social dimension must remain central. Otherwise, there is a risk of overlooking the very people and communities for whom these frameworks are intended to work, and to empower and protect.

In order to move from the theoretical definition to action, the GFC has developed a data equity framework to enable stakeholders to build more equitable data systems, processes and practices. Given that ethical and fairness issues relating to the use of data vary according to their specific context, the framework does not seek to be prescriptive or a “one-size-fits-all” solution. Rather, it is intended to prompt reflection, focus research and guide corrective action. Essentially, this framework should be regarded as a “framework for inquiry”, i.e. a guide to help spur conversations and evaluation inside organizations and communities as they consider using data, whether in AI-enabled systems or elsewhere. It is hoped that this framework will serve as a tool to uncover equity-related issues to be addressed within organizations.

The framework proposed here employs as its foundation the Te Mana o te Raraunga Model, an Indigenous model that describes the internal logic that traditional knowledge-keepers use when deciding to share knowledge with others. It considers data-sharing in relation to the nature of the data, the nature of data use, and the nature of the data user.<sup>9</sup> The Te Mana o te Raraunga Model informed Ngā Tikanga Paihere, a data ethics framework used to provide access to linked government data in New Zealand through an Integrated Data Infrastructure (IDI), and provides

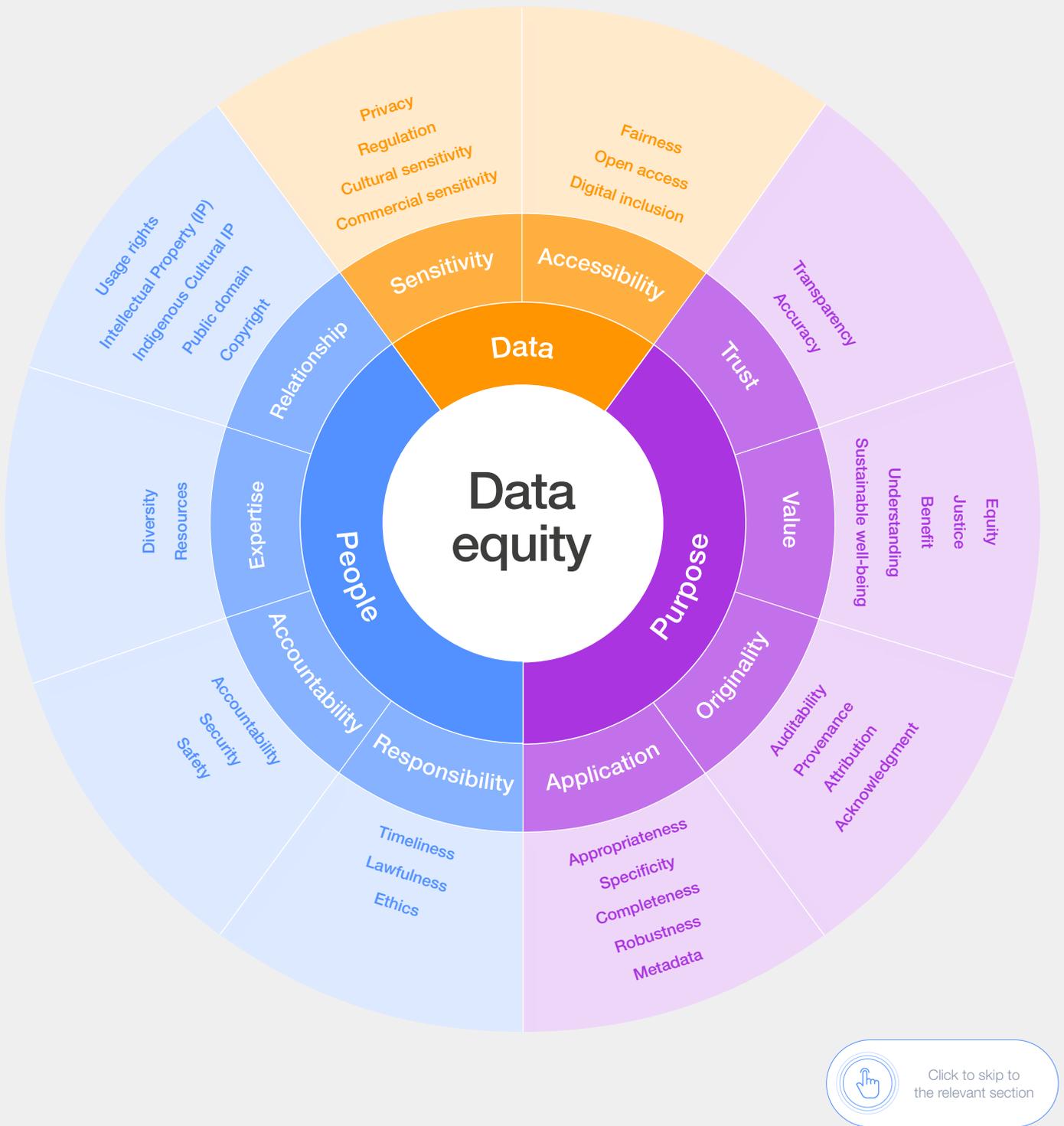
a useful lens for considering the broader issue of data equity.<sup>10</sup>

While developed in a unique cultural context, the Te Mana o te Raraunga Model has a broader applicability as it is aligned with the Five Safes Framework (safe data, projects, people, settings and outputs) that enables data services to provide safe research access to data.<sup>11</sup> The Five Safes Framework was adopted because of the central focus on human and social dimensions of equity, and consistency with the people and purpose-centric CARE Principles for Indigenous Data Governance (collective benefit, authority to control, responsibility and ethics).<sup>12</sup>

The CARE Principles are complementary to the data-centric FAIR Principles for Scientific Data Management (data should be findable, accessible, interoperable and reusable), which promote the use of open data.<sup>13</sup> The FAIR and CARE principles are promoted as key driving frameworks for data governance across a range of international and national policy environments (e.g. UNESCO Open Science Outlook,<sup>14</sup> AIATSIS Code of Ethics,<sup>15</sup> IPBES Data and Knowledge Management Policy<sup>16</sup> and World Data Systems Data Sharing Principles).<sup>17</sup> Similarly, the TRUST Principles (of transparency, responsibility, user focus, sustainability and technology) focus on the development of appropriate infrastructure for digital stewardship and preservation.<sup>18</sup>

Thus, building upon earlier work including the Te Mana o te Raraunga Model, the FAIR, CARE and TRUST Principles, the proposed data equity framework (Figure 1) is composed of 10 data equity characteristics grouped into three main categories: data, purpose and people. The **data** category is assessed in relation to its sensitivity and accessibility characteristics; the **purpose** category through its trust, value, originality and application characteristics; while the **people** category is associated with its relationship, expertise, accountability and responsibility characteristics.

FIGURE 1 | Data equity framework



Each of the 10 data equity characteristics in the framework is also associated with a set of related key issues, drawn from other existing data frameworks, including the OECD Privacy Principles,<sup>19</sup> EU Data Protection Principles,<sup>20</sup> OECD AI Principles,<sup>21</sup> Responsible Algorithm Principles,<sup>22</sup> Five Safes Framework,<sup>23</sup> FAIR Principles,<sup>24</sup> CARE Principles<sup>25</sup> and Indigenous Data Sovereignty Principles.<sup>26</sup> By mapping these issues to the framework’s core characteristics, seamless integration with existing data management principles is ensured. This approach not only aligns

with current best practices but also empowers practitioners and users to implement data equity seamlessly in their operations.

Data equity issues arise throughout the different phases of the data life cycle: during the **input stage** (collection and curation), the **process stage** (processing and analysis) and the **output stage** (visualization, sharing, application of resulting insights, benefit-sharing, reuse, retention and afterlife, and disposal), as displayed in Table 1.

TABLE 1 | Data equity considerations throughout the data life cycle (continued)

| Data life cycle stage   | Data life cycle sub-stage | Examples of data equity considerations at this stage   |
|---|---------------------------|--|
| <br>Input    | <b>Collection</b>         | <p>The collection of data is foundational for data equity. This stage is the most relevant intervention point to ensure proper data equity practices. Key questions to consider at this stage include:</p> <ul style="list-style-type: none"> <li>– Why is data collected? Is data collection necessary?</li> <li>– Who is collecting the data and how is the data collected?</li> <li>– Whose data is being collected and are they informed about (and possibly participating in the definition of) this collection?</li> <li>– Has the data been labelled? If so, have the labels been assessed for bias? Are solid methodological approaches used for labelling (e.g. multiple labellers, documented standards for labels, measurement of inter-rater reliability (IRR))?</li> </ul>  |
|   | <b>Curation</b>           | <p>Adherence to global principles and standards can inform curation of high-quality data.</p> <ul style="list-style-type: none"> <li>– Is the data representative of everyone that the system will impact?</li> <li>– Has the data been assessed for bias, toxicity and harmful representation?</li> <li>– Is the data accurate (because, for instance, data translated into other languages can embed bias and inaccuracies)?</li> </ul>  |
| <br>Process  | <b>Processing</b>         | <p>When processing data, the data pipeline should be transparent about how data is collected, processed and used.</p> <ul style="list-style-type: none"> <li>– Has the data been appropriately anonymized or pseudonymized?</li> <li>– Is the data being processed according to the purpose for which it was collected?</li> <li>– Is there a clear rationale for when data points are deemed “outliers” and deleted?</li> </ul>   |
|   | <b>Analysis</b>           | <p>Data collection, curation and processing directly informs analysis. Analysis may be skewed, discriminatory or flawed if the data used in the analysis is compromised or biased. Therefore, there has to be transparency, accuracy and fairness in the algorithms used for processing data.</p> <ul style="list-style-type: none"> <li>– Are the assumptions included in algorithms transparent, accurate, just and fair?</li> <li>– Have these been arrived at through participation of data subjects and impacted people/groups?</li> </ul>  |
| <br>Output | <b>Visualization</b>      | <p>Data visualization can reveal the power of data: it can mislead and undermine, as well as uplift and elucidate. The application of insights from the data visualization process is often key in shaping policy and outcomes.</p> <ul style="list-style-type: none"> <li>– Is the way the data has been visualized accurate and representative of, and in line with, the data?</li> <li>– Has data been visualized in a manner that clarifies it, rather than misrepresents it?</li> </ul>   |
|   | <b>Sharing</b>            | <p>Data sharing relates to fairness and transparency in how data is processed and shared. Alignment on global standards can facilitate data sharing.<sup>27</sup></p> <ul style="list-style-type: none"> <li>– Who has access to data and for how long?</li> <li>– How is data being shared, either across companies, agencies or across borders?</li> <li>– Do data subjects – individual and collective – have access to data? Is data shared with them?</li> <li>– What type of data is being shared globally and what is not? What legal frameworks/agreements protect data and data subjects’ interests in such cross-border sharing?</li> <li>– Is data shared in a safe and secure manner, adhering to applicable privacy regulations?</li> <li>– Are there limits on who can link data to other datasets?</li> <li>– Do data recipients have the necessary details about input and processing to make responsible decisions about data use, while respecting privacy?</li> </ul> |

TABLE 1 | Data equity considerations throughout the data life cycle (continued)

| Data life cycle stage   | Data life cycle sub-stage                          | Examples of data equity considerations at this stage  |
|---|--|---|
|  <p>Output</p> | <p><b>Application of resulting insights</b></p>    | <p>When using the insights generated from data to inform policies or regulations, it is essential to take into account prior considerations about the fairness of the data, including its collection and curation. Otherwise, the application of the resulting insights may result in biased or discriminatory policies.</p>  |
|   | <p><b>Benefit sharing<sup>28</sup></b></p>         | <p>Part of data equity is also considering who benefits from what data, and to what extent the people and communities whose data is collected also benefit from it.</p> <ul style="list-style-type: none"> <li>- Are benefits from data, whether in automated decision-making or otherwise, distributed equitably among individuals and communities?</li> <li>- Is the data subjects' right to benefit-sharing considered?</li> </ul>   |
|   | <p><b>Reuse</b></p>                                | <p>Similarly, when reusing previously collected data, earlier data equity considerations of collection and analysis resurface. Additionally, it is critical to assess whether previous datasets can effectively be reused in a different context.</p>   |
|   | <p><b>Retention and afterlife<sup>29</sup></b></p> | <p>Data retention refers to data storage and is an intrinsic aspect of data governance, linked to safety, security and privacy considerations. Additionally, storage capacity is a consideration for the longevity and legacy of the data.</p> <ul style="list-style-type: none"> <li>- What data is stored and for how long?</li> <li>- For what purpose is the data stored?</li> <li>- Is the data stored in a safe and secure manner, adhering to privacy and security considerations?</li> <li>- What data will perdure to inform policy and human outcomes?</li> <li>- How can you monitor for harm? How do you know if someone has been harmed because their data was hacked, was leaked to the public or used against them (for example, by being identified as "high-risk")?</li> </ul> |
|   | <p><b>Disposal</b></p>                             | <p>When thinking about ethical approaches to data science, a strong emphasis exists on the use and reuse of data. However, it is also important to consider what happens to data after it has been used.</p> <ul style="list-style-type: none"> <li>- In what ways is the data being disposed of and how can policies and regulations, such as the General Data Protection Regulation's (GDPR) right to erasure ("right to be forgotten"),<sup>30</sup> ensure equitable disposal of data?</li> <li>- Is harm potentially generated in deleting data?</li> </ul>  |

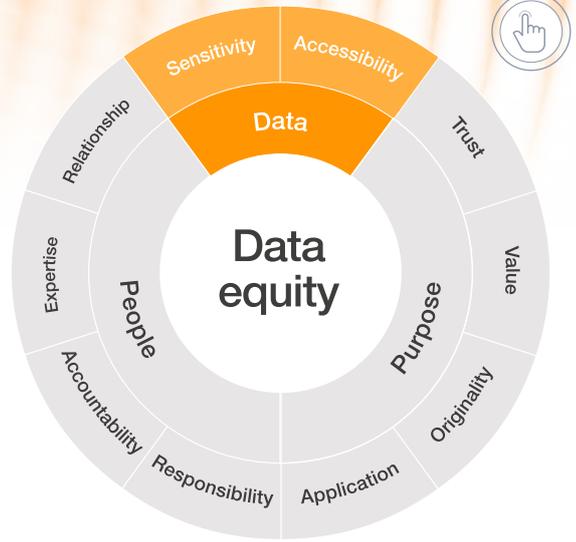
Given the impact of data on the digital society, it is crucial to ensure that data equity is preserved across all stages of the data life cycle, as quality and equity issues that are neglected in earlier stages cannot simply be remedied at a later stage. By considering data equity from the initial stage of data collection, inequitable practices and outcomes later in the data life cycle can be minimized. Moreover, data collectors should also consider

the possible subsequent (re)use of their data by other actors in potentially harmful ways.

Therefore, as part of the framework, a series of questions have been developed to consider data characteristics at the input, process and output stages, including a few suggested initial actions to implement data equity.<sup>31</sup> The questions and suggested actions are depicted below.



# Data



Examining various characteristics of data inputs into data analytics, including machine learning and genAI, can improve outcomes and ensure that biases are addressed early in the process.

## Sensitivity

### Key issues:

**Data harm potential:** What risks or negative impacts could result from the data's use?

**Privacy considerations:** How is personal information protected in the data?

**Regulation:** What legal frameworks govern the data's collection and use?

**Cultural sensitivity:** How does the data respect and impact diverse communities, cultural norms and values?

**Commercial sensitivity:** How is confidential business information safeguarded in the data?

### Suggested actions:

Review sensitive data requirements with privacy experts

Adopt transparent release strategies

Implement privacy and potential harm assessments

Ensure alignment of permissions for data access and re/use to Indigenous frameworks

Map all potential outcomes, and have negation/mitigation strategies for all possible negative outcomes

Map outcomes for all impacted stakeholder groups

## Accessibility

### Key issues:

**Fairness:** Does data collection, analysis and output lead to fair outcomes among impacted communities?

**Open access:** How accessible and transparent is the data, the algorithms used in data processing and the outputs of the data?

**Ability to share data:** How is data shared, in what manner, and who decides this?

**Interoperability:** Is data interoperable, to ensure accuracy, completeness and consistency in producing equitable outcomes?

**Digital inclusion:** How does data collection, analysis and output benefit all individuals and communities?

**Access regardless of one's abilities:** What processes can be put in place to ensure that everyone can access data?

**Access to data subjects (individual and collective):** How accessible is the data source, are data subjects (individual/collective) aware, and do they participate in data collection activities?

### Suggested actions:

Encourage alignment and participation

Develop open-code policies

Ensure data is accessible to individuals regardless of ability, especially to the data subjects concerned (individual/collective)

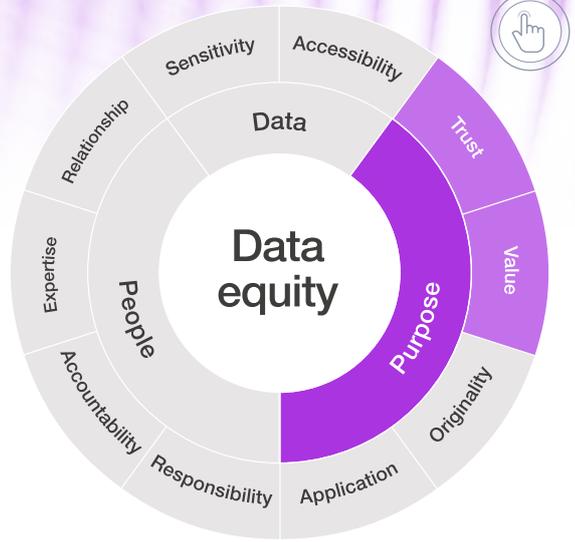
Ensure data is interoperable, through the use of harmonized standards where these exist

Be cognizant of data scarcity for under-represented communities and their languages

Ensure data of interest is findable, accessible and legible for relevant communities



# Purpose



Data analysis requires a clear purpose. Without one, analytics may lack fairness and impact, or even cause harm.

## Trust

### Key issues:

**Transparency:** How transparent are the data practices and policies?

**Bias:** What process is used to identify bias throughout the data life cycle?

**Explainability:** To what extent can data processes, including collection, analysis, output and policy decisions based on outputs be clearly explained?

**Accuracy:** What methods have been used to ensure quality, completeness and consistency?

**Control:** What methods are in place to ensure checks throughout the process?

### Suggested actions:

Make metadata available and understandable

Implement rigorous benchmarking against equitable datasets

Ensure that the training data is representative of the populations to be impacted by the system

Embed model and system traceability and accountability

Disclose non-human interaction

Make disclosures to Indigenous communities about Indigenous data

Continuously monitor for harm

## Value

### Key issues:

**Human rights:** Does data collection, analysis and output respect and promote human rights?

**Justice:** Is the value of data considered in a fair and just manner?

**Benefit-sharing:** Who benefits from the value generated by the data and how are these benefits distributed?

**Understanding:** Are cultural and social norms understood and have communities been consulted in data usage?

**Sustainable well-being:** Does the output generated by data contribute to long-term social and environmental well-being?

**Value for whom (individual and collective):** Who decides the value of data and who obtains it?

### Suggested actions:

Focus on human values and preferences

Build public awareness of AI capabilities and their limitations

Ensure a role in value determination and accrual for data subjects – both individual and communities

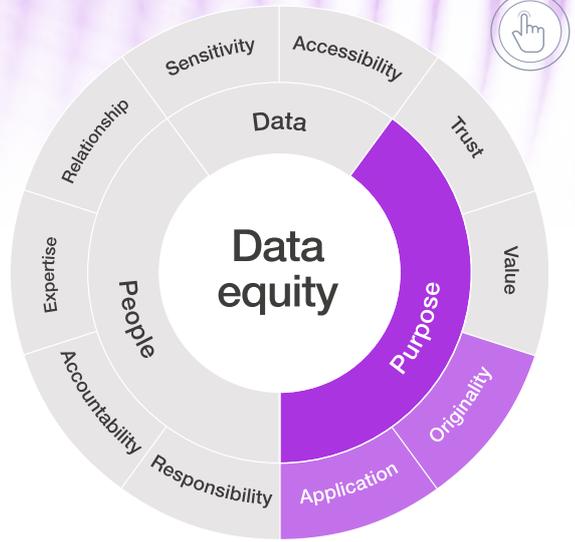
Ensure Indigenous peoples and other vulnerable groups determine the benefits of their data

Build relationships with Indigenous peoples and other vulnerable groups when collecting their data

Collect data relevant to Indigenous languages and worldviews with consent and in a culturally appropriate manner



# Purpose



Data analysis requires a clear purpose. Without one, analytics may lack fairness and impact, or even cause harm.

## Application

### Key issues:

**Appropriateness:** Is the data suitable for its intended purpose?

**Accuracy:** Is the data used accurately and assessed to be consistent for its purpose?

**Specificity:** Is the data specific enough for the intended purpose?

**Representativeness:** Does the dataset represent its specific purpose and the populations that will be affected by the results, both at the individual and community level?

**Robustness:** Is the dataset of sufficient size and scale?

**Metadata availability:** Is the metadata complete, fit for purpose and accessible?

### Suggested actions:

Adopt sandbox processes

Develop comprehensive multi-level measurement frameworks

Indicate the representativeness of the data

Utilize Indigenous and culturally specific identifiers

Conduct data needs assessments

Enable culturally specific metadata fields

**Note:** A sandbox is an environment where technologies, services and business models can be tested in the market with real consumers. Regulatory requirements are relaxed or made flexible, often for a limited time period, but with appropriate supervision and safeguards.

## Originality

### Key issues:

**Auditability:** Has documentation been maintained to ensure that the analytical process can be audited and/or reviewed?

**Provenance:** Can the origin, journey and usage rights of the data be traced?

**Attribution:** Is attribution to the source data and contributors necessary?

**Acknowledgement:** Are source datasets and contributors recognized in the outputs?

**Derivativeness of work:** Are the data sources unique or is the data used for a new purpose?

### Suggested actions:

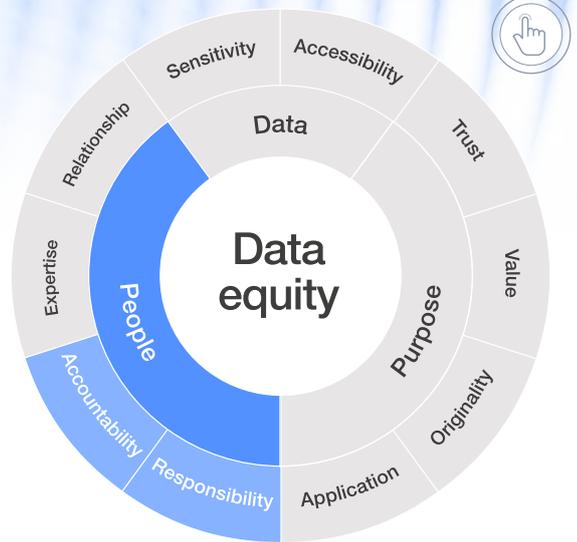
Ensure content traceability

Establish precise and shared terminology (including culturally specific metadata)

Promote equitable attribution, including acknowledgment and authorship



# People



Protecting individuals' data rights throughout the data life cycle is crucial to ensure that the collection and use of data benefit people and communities.

## Accountability

### Key issues:

**Security:** How is the data protected from unauthorized access, use or breaches?

**Safety:** What protocols are in place to prevent harm from data use?

**Auditability:** Has clear documentation been maintained of the development process and the related governance decisions?

**Control:** Who has decision-making power over the data and how it is used?

**Ownership:** What assessment of data ownership rights has been completed?

**Authority:** Under what authority has information been collected, used, shared and stored?

**Usage rights:** How have usage rights been obtained and documented?

**Access rights:** Is there a method for individuals to access their personal and sensitive information?

**Benefit rights:** Is benefit-sharing ensured for all stakeholders?

**Purpose limitation:** Are there restrictions on how the data can be used and reused?

**Participation:** How are stakeholders, individuals and communities involved in data-related decisions?

### Suggested actions:

Develop frameworks for data rights, ownership rights and benefit-sharing for data subjects (individuals and communities)

Develop contextual ways of implementing and auditing compliance with these frameworks

Enable user feedback and audit of people's data

Ensure communities' approval of outputs

## Responsibility

### Key issues:

**Timeliness:** Are there controls to ensure that data remains current and updated regularly?

**Lawfulness:** What laws, regulations and standards govern the type of data being used?

**Ethics:** What ethical considerations, which may harm individuals or the community, should be taken into account in data practices?

**Harmonization:** How will conflicts be managed and data practices harmonized across different contexts?

**Global standards:** When designing for multiple locations, what global standards will be used, and how will variations in requirements be handled?

### Suggested actions:

Implement ethical impact assessments

Implement step-by-step review

Ensure transparent ethics approval processes

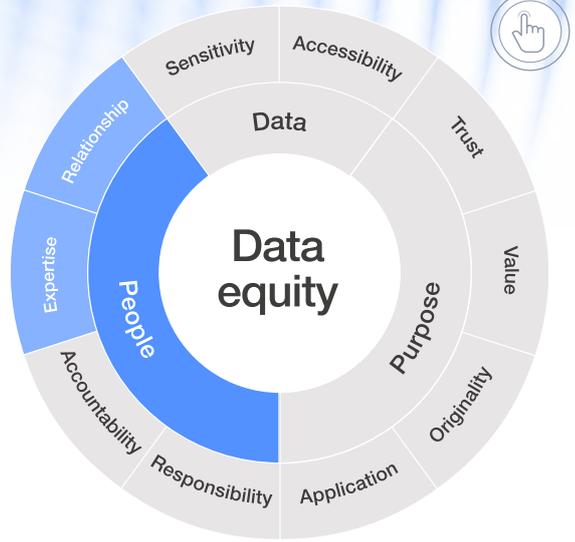
Ensure transparent processes to obtain community permissions

Implement safeguards to protect the mental well-being of individuals labelling data, particularly if the data is harmful

Pay a living wage to community members for their time and expertise



# People



Protecting individuals’ data rights throughout the data life cycle is crucial to ensure that the collection and use of data benefit people and communities.

## Expertise

### Key issues:

**Diversity:** How well does the data team represent different groups and perspectives, and have they received proper diversity, equity and inclusion training?

**Resources:** What specialized expertise is needed?

**Sociocultural expertise:** Have members of the cultures and societies affected been consulted?

### Suggested actions:

Employ diverse teams across the process including red teams

Fund training and education

Support community capacity-building

Ensure impacted communities are part of outcome assessments

## Relationship

### Key issues:

**Usage rights:** Who has the right to use the data and how?

**Access rights:** Who can view, access or obtain the data and who decides this?

**Benefit rights (individual and collective):** Will the outcomes be beneficial to the impacted individuals and/or communities?

**Intellectual property (IP):** What intellectual property protections need to be considered in using the data, or in generating new insights from the information?

**Indigenous cultural intellectual property (ICIP):** How are Indigenous knowledge and cultural expressions protected in the data?

**Public domain:** If data is obtained in the public domain, what recognition is required of source/prior use?

**Purpose:** Is the data being used for the purpose as originally designed?

### Suggested actions:

Adapt to the evolving landscape of creativity and IP

Develop frameworks of benefit-sharing with data subjects (individuals and communities), and means of the actual framework implementation

Adopt strategies to recognize ICIP

Ensure recognition of data sovereignty and Indigenous data sovereignty

Ensure recognition of Indigenous peoples’ and other communities’ rights to FPIC (free, prior and informed consent)

## 3 Case studies

These case studies demonstrate the data equity framework through real-world examples that can be adapted to other contexts.

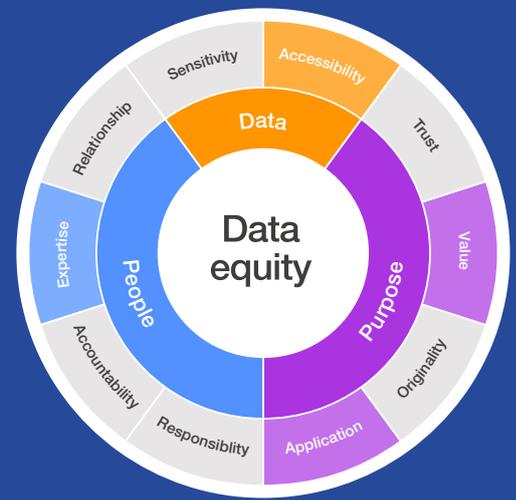
Each case study highlights specific data equity characteristics and issues that may arise at various stages of the data life cycle. The framework is not a linear roadmap but a flexible and iterative tool for critical reflection and inquiry, empowering users

to identify and address data equity concerns in their unique contexts. The proposed actions are intended to be a starting point, to awaken creativity and not to limit the possibilities in addressing the challenges identified.



# Climate data collection and monitoring

🔍 Robust climate data for mitigation strategies



## 📖 Context

Climate data collection and monitoring in developing countries are crucial for effective mitigation and adaptation strategies for addressing the impacts of climate change, in alignment with the United Nations Sustainable Development Goal 13.<sup>32</sup> Yet, significant gaps exist in climate data collection, especially in rural and remote areas.



## ➔ Input

### Data equity issues

- Significant gaps remain in climate data collection, particularly in rural and remote areas in the Global South, which hampers a comprehensive understanding of localized climate impacts.

### Suggested actions (not exhaustive)

- Invest in data collection to improve the granularity of climate data, especially for vulnerable communities.
- Invest in (community) capacity-building to enable and incentivize more effective climate monitoring.

## 🔄 Process

### Data equity issues

- There is a shortage of experts capable of translating technical climate data into culturally relevant information.

### Suggested actions (not exhaustive)

- Create or support regional multistakeholder climate monitoring networks with shared resources, best practices, and harmonized definitions and data-processing standards.

## ➔ Output

### Data equity issues

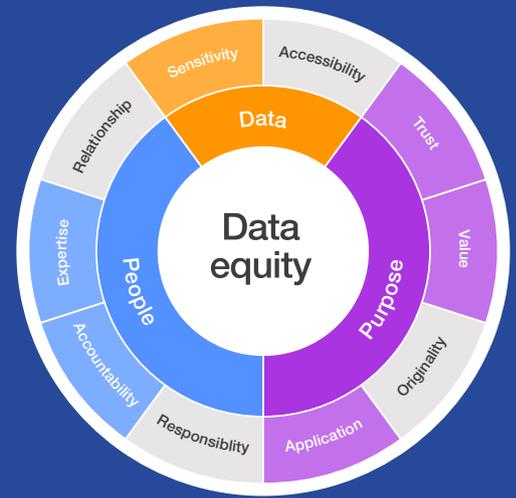
- Climate action strategies may be based on incomplete or non-representative data, potentially leading to inefficient or missed opportunities for climate change mitigation.

### Suggested actions (not exhaustive)

- Share climate data in culturally appropriate formats, including the use of local languages and storytelling techniques.
- Fund training and education of local researchers and decision-makers.
- Disclose possible limitations to inform end users.

# Women's access to financial services in emerging markets

🔍 Innovative data solutions to empower female entrepreneurs



## 📖 Context

In emerging markets, female entrepreneurs face significant challenges in accessing financial services, due to gender-biased lending practices. These systemic issues are due to a lack of data, credit systems that exacerbate gender disparities and reliance on biased data sources and analytical methods. Addressing this inequality requires innovative approaches to data, data analytics and algorithmic development to create equitable outcomes that reflect today's society.<sup>33</sup>



## ➔ Input

### Data equity issues

- Gender disaggregated data is not always available.
- Women's financial contributions may be part of family assets.
- Data on informal economy may not be included.

### Suggested actions (not exhaustive)

- Increase women's representation in the data.
- Create equitable synthetic/proxy data where data is not available.

## ⚙️ Process

### Data equity issues

- Algorithms are created using traditional methods that do not correct for inequities.
- Even when men and women have the same credit score, women are disproportionately rejected for loans.<sup>34</sup>

### Suggested actions (not exhaustive)

- Check credit algorithms for overt and covert bias, including proxy discrimination.
- Ensure even application of algorithms to both male and female datasets and adjust accordingly.

## ➔ Output

### Data equity issues

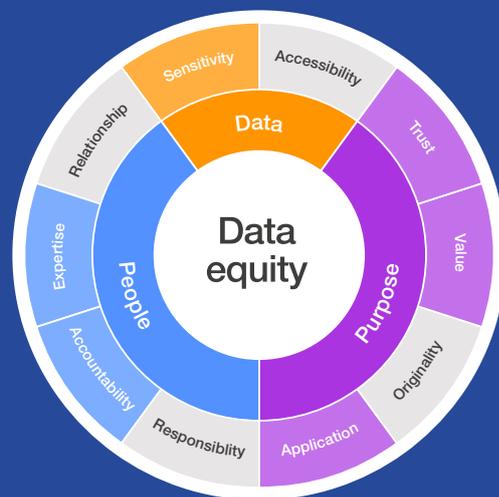
- Low representation in existing datasets is perpetuated by genAI models.

### Suggested actions (not exhaustive)

- Regularly assess algorithm performance to eliminate gender biases.
- Improve outcomes by ensuring female-owned businesses receive a percentage of loans.

# Racial bias in healthcare

🔍 Algorithm auditing to improve access to healthcare<sup>35</sup>



## Context

Racial bias in healthcare can lead to disparities in treatment and outcomes for patients. Commercial algorithms used to identify patients for complex care have historically disadvantaged Black patients compared to White patients, by predicting healthcare costs as opposed to illness severity. Addressing this bias could increase Black patients' access to care from 17.7% to 46.5%, underscoring the need for rigorous algorithm auditing and cross-sector collaboration to eliminate such biases in decision-making.<sup>36</sup>



## Input

### Data equity issues

- Historical data on healthcare costs used in algorithms reflects existing racial disparities.
- While race is explicitly excluded as an input variable, other variables correlating with race can lead to proxy discrimination.

### Suggested actions (not exhaustive)

- Collect more comprehensive health data, including direct measures of health status and barriers to healthcare access.
- Carefully audit input variables for potential proxy discrimination.



## Process

### Data equity issues

- Predicting future healthcare costs as a proxy for health needs disadvantages Black patients, who have historically not received expensive treatments.

### Suggested actions (not exhaustive)

- Maintain transparency in data collection and algorithmic scoring processes.



## Output

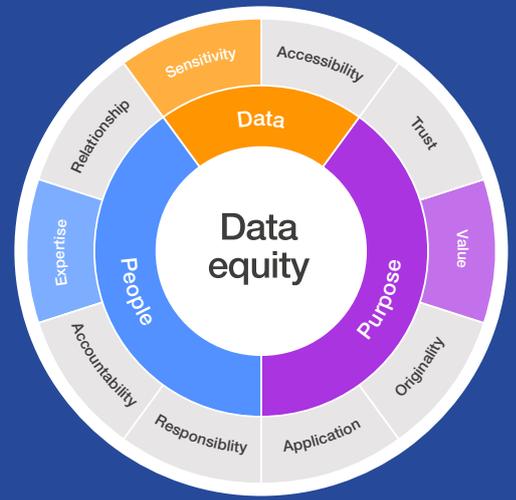
### Data equity issues

- The biased algorithmic output influences the human decision-making of physicians, who only partially mitigate the algorithmic bias.

### Suggested actions (not exhaustive)

- Regularly audit the impact of algorithmic decisions on patient outcomes across different racial groups.
- Empower clinicians to flag potentially biased or incorrect predictions.

# Improving how the City of Boston collects gender data



🔗 Development of gender identity guidelines with the LGBTQ+ community in Boston

## 📖 Context

When asking residents about gender identity to deliver key services, governments rarely use gender inclusive language. Forcing gender binaries can lead to data collection that is misrepresentative of people’s gender identities. Additionally, collecting this sensitive information can increase risk of harm and barriers to participation for vulnerable minorities.

Therefore, the City of Boston partnered with members of the local LGBTQ+ community (lesbian, gay, bisexual, transgender, queer or questioning persons and others) to develop guidelines for how city officials should collect data about gender identity.<sup>37</sup>



## ➔ Input

### Data equity issues

- Collection of gender identity data in a binary manner can lead to bias and misrepresentation of communities.

### Suggested actions (not exhaustive)

- Establish clear criteria for when to collect gender identity data.
- Incorporate flexible data collection methods to ensure privacy and autonomy, and implement de-gendered processes wherever appropriate.

## ⚙️ Process

### Data equity issues

- Processing issues may arise if wider systems enforce binary options or fail to effectively account for non-binary identities.
- Discrepancies between city, state and federal data systems can lead to inter-jurisdictional issues.

### Suggested actions (not exhaustive)

- Provide affirming, respectful guidelines for asking about gender identity, including multiple response options and privacy-focused data collection mechanisms.
- Be cognizant of data scarcity for under-represented or marginalized communities.

## ➔ Output

### Data equity issues

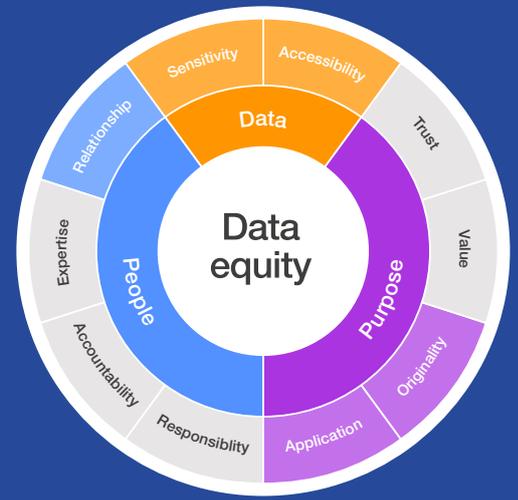
- Service outcomes may be biased if the data collected inadequately represents gender identities of vulnerable minorities.

### Suggested actions (not exhaustive)

- Ensure compliance with gender identity data standards across all city services and systems, including third-party data management.
- Integrate robust data security policies to protect gender identity data throughout its life cycle.

# Mapping Indigenous territories

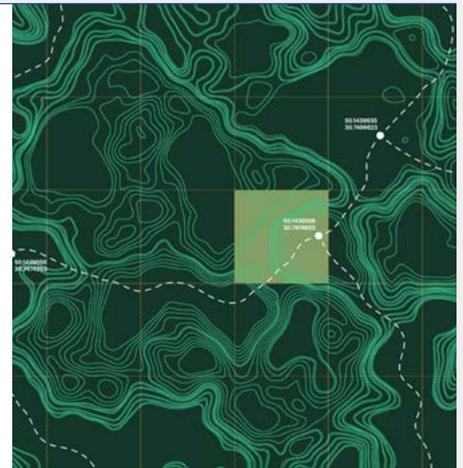
📍 Recognition of Indigenous territories on maps to ensure equitable representation



## 📖 Context

Inadequate representation of Indigenous territories on digital mapping platforms endangers cultural identities by ignoring ancestral landmarks and boundaries, thus limiting access to basic services and perpetuating marginalization. Accurate mapping is essential for documenting land claims, supporting environmental planning and ensuring emergency preparedness.<sup>38</sup>

As a result of a seven-year collaboration between Canadian Indigenous communities and Google Earth, Indigenous lands are now recognized on Google Maps. The same initiative has also brought visibility to Indigenous territories in Brazil, where users can now observe the conservation efforts of different ethnic groups in the Amazon.<sup>39</sup>



## ➔ Input

### Data equity issues

- Issues persist in terms of representation of Indigenous lands on mapping services, including consent to representation and ownership of geographical data.

### Suggested actions (not exhaustive)

- Consult with Indigenous communities to ensure that data collection respects Indigenous sovereignty and self-determination.
- Obtain informed consent and clearly define how data will be used and represented.

## ⚙️ Process

### Data equity issues

- Algorithms may inadvertently prioritize certain geographical features or landmarks over others, while insufficient cultural sensitivity protocols may lead to culturally offensive output.

### Suggested actions (not exhaustive)

- Integrate cultural sensitivity training and diverse perspectives in algorithm development and data processing teams.
- Promote equitable attribution to Indigenous data sources.

## ➔ Output

### Data equity issues

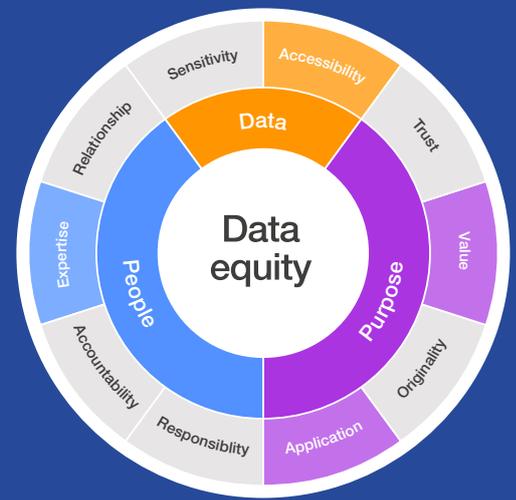
- Misrepresentation of Indigenous lands can perpetuate cultural erasure.
- Incorrect mapping can affect Indigenous sovereignty and land rights, and influence legal decisions.

### Suggested actions (not exhaustive)

- Provide mechanisms for Indigenous communities to review and verify mapped data, and options to control the visibility of certain locations or sensitive information on public maps.

# National Statistical Offices as data stewards

🔍 Ensuring equitable data stewardship for informed policy-making



## Context

National Statistical Offices (NSOs) are evolving into data stewards, integrating diverse data systems. Official statistics' adherence to recognized principles ensures that they remain a trusted, freely accessible public resource. As primary data handlers and producers, NSOs must prioritize equity, as their data shapes policies and initiatives in the country and across sectors.<sup>40</sup>



## Input

### Data equity issues

- Surveys often falter when people distrust the process, find it demanding or fail to see personal value, hampering crucial data collection efforts.

### Suggested actions (not exhaustive)

- Ensure sufficient resources are available for comprehensive and representative data collection.
- Reduce burden by implementing the “ask once” principle and promoting interoperability across data sources.



## Process

### Data equity issues

- A lack of consistent definition and standards may result in skewed results or limit the usefulness of the data.

### Suggested actions (not exhaustive)

- Implement statistical capacity-building programmes while ensuring the use of harmonized standards.
- Ensure analysis provides relevant insights for specific communities.



## Output

### Data equity issues

- Misinterpretation of the data and methodologies, and the lack of harmonization and comparability with published results can limit the utility of the output.

### Suggested actions (not exhaustive)

- Ensure equitable value creation from the data collected.
- Share metadata and methodology in an accessible and transparent manner, ensuring the use of privacy enhancing technologies (PETs).

4

# Recommendations for stakeholders

This section offers an initial set of suggested actions to guide key stakeholders in addressing data equity issues.

Ensuring fair and equitable outcomes for all through responsible use of data is a collective duty. While the challenges vary by context, the following table summarizes some proposed actions that key actors – from data collectors and regulators to end users – should take into account in developing strategies to address the different characteristics of data equity. While not

an exhaustive list, these recommendations are based on the proposed data equity framework and provide a general map of issues that stakeholders should prioritize. However, it is important to note that many of these issues are common to multiple stakeholders and would benefit from collaboration among them for more effective implementation.

TABLE 2 Recommendations for key stakeholders to implement data equity

|  |   |
|--|---|
|  <p><b>Private-sector companies</b></p>        | <ul style="list-style-type: none"> <li>- Adapt to the evolving landscape of creativity and IP</li> <li>- Adopt transparent ethics approval processes</li> <li>- Adopt transparent release strategies</li> <li>- Disclose non-human interaction</li> <li>- Embed model and system traceability and accountability</li> <li>- Employ diverse red teams</li> <li>- Enable user feedback and audit of people’s data</li> <li>- Implement ethical impact assessments</li> <li>- Implement rigorous benchmarking against equitable datasets</li> <li>- Implement transparent and inclusive auditing mechanisms</li> </ul>                     |
|  <p><b>Academia and technical experts</b></p> | <ul style="list-style-type: none"> <li>- Collect data relevant to Indigenous languages and worldviews with consent and in a culturally appropriate manner</li> <li>- Develop comprehensive multi-level measurement frameworks</li> <li>- Establish precise and shared terminology (including culturally specific metadata)</li> <li>- Implement rigorous benchmarking against equitable datasets</li> <li>- Perform ethical impact assessments</li> <li>- Promote equitable attribution, including acknowledgement and authorship</li> <li>- Provide training and educational programmes</li> </ul>                                     |
|  <p><b>Government/ public sector</b></p>      | <ul style="list-style-type: none"> <li>- Adopt sandbox processes</li> <li>- Develop open-code policies</li> <li>- Disclose non-human interaction</li> <li>- Ensure recognition of Indigenous data sovereignty and Indigenous peoples’ rights to free, prior and informed consent</li> <li>- Ensure transparent community permissions processes</li> <li>- Fund training and education, and support community capacity-building</li> <li>- Harmonize standards internationally, while respecting regional norms</li> <li>- Harmonize standards for data input, processing and output</li> <li>- Implement privacy assessments</li> </ul> |

TABLE 2 | Recommendations for key stakeholders to implement data equity (continued)

|  |  |
|--|--|
|  <p><b>National Statistical Offices</b></p>       | <ul style="list-style-type: none"> <li>– Adopt transparent processes for obtaining ethics approval for data collection, processing and dissemination</li> <li>– Collect data relevant to Indigenous languages and worldviews with consent and in a culturally appropriate manner</li> <li>– Conduct data needs assessment</li> <li>– Enable culturally specific metadata fields</li> <li>– Ensure data of interest is findable, accessible and legible for relevant communities</li> <li>– Ensure transparent and inclusive processes for obtaining community permissions</li> <li>– Utilize Indigenous and culturally specific identifiers</li> </ul>                     |
|  <p><b>Civil society organizations (CSOs)</b></p> | <ul style="list-style-type: none"> <li>– Build public awareness of AI capabilities and their limitations</li> <li>– Build relationships with Indigenous peoples and other vulnerable groups and adopt strategies to recognize Indigenous cultural and intellectual property (ICIP)</li> <li>– Conduct data needs assessments</li> <li>– Encourage transparency, privacy assessments and alignment of permissions for data access</li> <li>– Focus on human values and preferences</li> <li>– Perform ethical impact assessments</li> <li>– Promote equitable attribution including acknowledgment and authorship</li> <li>– Support community capacity-building</li> </ul> |
|  <p><b>General public</b></p>                    | <ul style="list-style-type: none"> <li>– Encourage alignment and participation, including in community capacity-building and education</li> <li>– Encourage transparency, privacy assessments and alignment of permissions for data access</li> </ul>  |
|  <p><b>Communities</b></p>                      | <ul style="list-style-type: none"> <li>– Adapt to the evolving landscape of creativity and IP</li> <li>– Contribute to data needs assessment</li> <li>– Promote alignment of permissions for data access and re/use to Indigenous frameworks</li> <li>– Promote Indigenous and minority groups' approval of outputs</li> <li>– Promote transparent processes for obtaining community permissions</li> </ul>  |

# Conclusion

Stakeholders are encouraged to champion and integrate these principles in their operations and decision-making processes.

The essence of data equity transcends technical processes; it is fundamentally about the impact on people and communities. Thus, as technical capabilities advance, it is imperative that the awareness of their social implications does too.

In the pursuit of a more equitable world, the data equity definition and framework introduced in this report seek to serve not merely as a set of guidelines but as dynamic tools, urging all stakeholders across sectors involved in the realms of data and technology to prioritize and operationalize equity at every stage of their work.

Implementing the proposed data equity framework from the onset of any data-related initiative is crucial. The iterative and adaptable nature of the framework seeks to spark ongoing dialogue and continuous improvement in data practices and encourage stakeholders to consistently assess how data practices affect diverse groups.

Stakeholders are asked to not simply adopt this framework, but to champion and integrate its principles into the fabric of their operations and decision-making processes. By embedding these considerations into discussions at all levels – from product development to strategic leadership – organizations can begin to assess their current practices and identify crucial areas for improvement.

The Global Future Council on Data Equity is dedicated to forging a future where cutting-edge technologies empower all, and to ensuring that fairness and inclusivity drive both technological advancements and their real-world applications. The framework introduced here is designed to be a crucial foundation for transforming data practices to fully embrace inclusivity and fairness. By achieving this, the aim is to ensure that the era of digital transformation is characterized not only by technological breakthroughs, but also by significant social advancements.

# Contributors

## Global Future Council on Data Equity 2023-2024

The World Economic Forum's Network of Global Future Councils is the world's foremost multistakeholder and interdisciplinary knowledge network dedicated to promoting innovative thinking to shape a more resilient, inclusive and sustainable future.

## Global Future Council on Data Equity members

### Co-chairs

#### JoAnn Stonier

Mastercard Fellow, Data and Artificial Intelligence, Mastercard

#### Lauren Woodman

Chief Executive Officer, DataKind

### Members

#### Majed Alshammari

Special Adviser, Data Governance, Saudi Data and AI Authority (SDAIA)

#### Kathy Baxter

Principal Architect, Responsible Artificial Intelligence and Technology, Salesforce

#### Alberto Giovanni Busetto

Chief Artificial Intelligence Officer, HealthAI

#### Renée Cummings

Data Science Professor and Data Activist in Residence, University of Virginia

#### Nighat Dad

Founder and Executive Director, Digital Rights Foundation

#### Arti Garg

Head of Technology Strategy and Evaluation and Senior Distinguished Technologist, Hewlett Packard Enterprise

#### Katherine Hsiao

Executive Vice-President; Head, Health and Life Sciences, Palantir Technologies

#### Maui Hudson

Associate Professor and Director, Te Kotahi Research Institute, University of Waikato

#### David Kanamugire

Chief Executive Officer, National Cyber Security Agency of Rwanda

#### Astha Kapoor

Co-Founder, Aapti Institute

#### Zheng Lei

Professor, Fudan University

#### María Paz Canales Loebel

Head of Legal, Policy and Research, Global Partners Digital

#### Jacqueline Lu

President and Co-Founder, Helpful Places (DTPR)

#### Angela Oduor Lungati

Executive Director, Ushahidi

#### Emna Mizouni

Chief Executive Officer, Digital Citizenship

#### Parminder Jeet Singh

Digital Society Researcher

#### Sarah Telford

Lead, Centre for Humanitarian Data, United Nations Office for the Coordination of Humanitarian Affairs (OCHA)

#### Georges-Simon Ulrich

Director General, Swiss Federal Statistical Office (FSO)

## World Economic Forum

### Council managers

#### Karla Yee Amezaga

Lead, Data Policy and AI, Centre for the Fourth Industrial Revolution

#### Stephanie Teeuwen

Specialist, Data Policy and AI, Centre for the Fourth Industrial Revolution

# Acknowledgements

## **Genta Ando**

Executive Director, Japan External Trade Organization; Fellow, Centre for the Fourth Industrial Revolution

## **Kimmy Bettinger**

Lead, Expert and Knowledge Communities, Centre for the Fourth Industrial Revolution

## **Daniel Dobrygowski**

Head, Governance and Trust, Centre for the Fourth Industrial Revolution

## **Daisuke Fukui**

Senior Researcher, Hitachi America; Fellow, Centre for the Fourth Industrial Revolution

## **Rafi Lazerson**

GenAI Policy Manager, Accenture; Fellow, Centre for the Fourth Industrial Revolution

## **Cathy Li**

Head, AI, Data and Metaverse, Centre for the Fourth Industrial Revolution; Member of the Executive Committee

## **Dylan Reim**

Lead, Metaverse Governance, Centre for the Fourth Industrial Revolution

## **Hannah Rosenfeld**

Specialist, Artificial Intelligence and Machine Learning, Centre for the Fourth Industrial Revolution

## **Stephanie Smittkamp**

Coordinator, Artificial Intelligence and Data Team, Centre for the Fourth Industrial Revolution

### **Additional acknowledgements**

## **Alejandro Jimenez Jaramillo**

Director of Governance and Policy, City of Boston

## **Manakore Rickus-Graham**

Kaiatawhai Raraunga Māori, Nicholson Consulting

## **Production**

### **Laurence Denmark**

Creative Director, Studio Miko

### **Madhur Singh**

Editor, World Economic Forum

### **Oliver Turner**

Designer, Studio Miko

# Endnotes

1. World Economic Forum. Global Future Council on the Future of Data Equity. <https://www.weforum.org/communities/gfc-on-data-equity/>.
2. World Economic Forum. (2023). *Data Equity: Foundational Concepts for Generative AI*. <https://www.weforum.org/publications/data-equity-foundational-concepts-for-generative-ai/>.
3. World Economic Forum. (2023). *Data Equity: Foundational Concepts for Generative AI*. <https://www.weforum.org/publications/data-equity-foundational-concepts-for-generative-ai/>.
4. Some academic sources on the concept of data equity include: Morey, B.N. et al. (2022). No Equity without Data Equity: Data Reporting Gaps for Native Hawaiians and Pacific Islanders as Structural Racism. *Journal of Health Politics, Policy and Law*, vol. 47, no. 2, 1 April 2022, pp. 159-200. <https://doi.org/10.1215/03616878-95171777>; Gee, G.C. et al. (2022). Considerations of Racism and Data Equity Among Asian Americans, Native Hawaiians, and Pacific Islanders in the Context of COVID-19. *Social Epidemiology*, vol. 9, 2022, pp. 77-86. <https://link.springer.com/article/10.1007/s40471-022-00283-y>; Prainsack, B. et al. (2022). Data Solidarity, *Governing Health Futures 2030*. <https://www.governinghealthfutures2030.org/wp-content/uploads/2022/12/DataSolidarity.pdf>; Buolamwini, J. (2023). *Unmasking AI: My Mission to Protect What is Human in a World of Machines*. Random House. Mejias, U.A., & Couldry, N. (2024). *Data Grab: The New Colonialism of Big Tech and How to Fight Back*. The University of Chicago Press.
5. An earlier publication of the Global Future Council on Data Equity considers these four classes of data equity and acknowledges how these are influenced and impacted by equitable practices and considerations in procedures and decision-making: Jagadish, H., Stoyanovich, J., & Howe, B. (2023). The Many Facets of Data Equity, *Journal of Data and Information Quality*, vol. 14, no. 4, February 2023. <https://doi.org/10.1145/3533425>; World Economic Forum. (2023). *Data Equity: Foundational Concepts for Generative AI*, pp. 4-5. <https://www.weforum.org/publications/data-equity-foundational-concepts-for-generative-ai/>.
6. World Economic Forum. (2024). *Data Equity Definition*. [https://www3.weforum.org/docs/WEF\\_Data\\_Equity\\_Definition\\_2024.pdf](https://www3.weforum.org/docs/WEF_Data_Equity_Definition_2024.pdf).
7. World Economic Forum. (2022). *Earning Digital Trust: Decision-Making for Trustworthy Technologies*. <https://www.weforum.org/publications/earning-digital-trust-decision-making-for-trustworthy-technologies/>.
8. World Economic Forum. (2023). *Data Equity: Foundational Concepts for Generative AI*, pp. 4-5. <https://www.weforum.org/publications/data-equity-foundational-concepts-for-generative-ai/>.
9. Hudson, M. et al. (2017). He matapihi ki te Mana Raraunga – Conceptualising Big Data through a Māori lens. In H. Whaanga, T. T. Keegan, & M. Apperley (Eds.), *He Whare Hangarau Māori - Language, culture & technology* (pp. 64-73). Te Whare Wānanga o Waikato (University of Waikato, Kirikiriroa; Hamilton, New Zealand); Te Pua Wānanga ki te Ao (Faculty of Māori and Indigenous Studies). [https://www.academia.edu/108801362/He\\_Matapihi\\_ki\\_te\\_Mana\\_Raraunga\\_Conceptualising\\_Big\\_Data\\_through\\_a\\_M%C4%81ori\\_lens](https://www.academia.edu/108801362/He_Matapihi_ki_te_Mana_Raraunga_Conceptualising_Big_Data_through_a_M%C4%81ori_lens).
10. Stats NZ (2020). *Ngā Tikanga Paihere: a framework guiding ethical and culturally appropriate data use*. <https://data.govt.nz/assets/data-ethics/Nga-Tikanga/Nga-Tikanga-Paihere-Guidelines-December-2020.pdf>.
11. UK Data Service. *What is the Five Safes framework*, SecureLab. <https://ukdataservice.ac.uk/help/secure-lab/what-is-the-five-safes-framework/>.
12. Global Indigenous Data Alliance (GIDA) (2018). *CARE Principles for Indigenous Data Governance*. <https://www.gida-global.org/care>; Carroll, S. R. et al. (2020). The CARE Principles for Indigenous Data Governance. *Data Science Journal*, vol. 19, no.1. <https://doi.org/10.5334/dsj-2020-043>.
13. Wilkinson, M.D. et al. (2016, Mar 15). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, vol. 3, no. 160018, Mar 15, 2016, pp. 1-9. <https://doi.org/10.1038/sdata.2016.18>.
14. UNESCO. (2023). *Open science outlook 1: status and trends around the world*. <https://unesdoc.unesco.org/ark:/48223/pf0000387324>.
15. Australian Institute of Aboriginal and Torres Strait Islander Studies (AIATSIS). (2020). *AIATSIS Code of Ethics for Aboriginal and Torres Strait Islander Research*. <https://aiatsis.gov.au/sites/default/files/2020-10/aiatsis-code-ethics.pdf>.
16. Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES). (2022). *Data and knowledge management policy*. <https://www.ipbes.net/node/38260>.
17. World Data System. *Data Sharing Principles*. <https://worlddatasystem.org/about/data-sharing-principles/>.
18. Lin, D. et al. (2020, May 14). The TRUST Principles for digital repositories. *Scientific data*, vol. 7, no. 144, May 14, 2020. <https://www.nature.com/articles/s41597-020-0486-7>.
19. OECD. (2013). *The OECD Privacy Framework*. <https://www.oecd.org/en/topics/policy-issues/privacy-and-data-protection.html>.
20. European Commission. (2016). *Data protection in the EU*. [https://commission.europa.eu/law/law-topic/data-protection/data-protection-eu\\_en](https://commission.europa.eu/law/law-topic/data-protection/data-protection-eu_en).
21. OECD. (2019). *OECD AI Principles overview*, OECD AI Policy Observatory. <https://oecd.ai/en/ai-principles>.
22. OECD. (2022). *Principles for Accountable Algorithms and a Social Impact Statement for Algorithms*, OECD.AI Policy Observatory. <https://oecd.ai/en/catalogue/tools/principles-for-accountable-algorithms-and-a-social-impact-statement-for-algorithms>.

23. UK Data Service. *What is the Five Safes framework*, SecureLab. <https://ukdataservice.ac.uk/help/secure-lab/what-is-the-five-safes-framework/>.
24. Wilkinson, M.D. et al. (2016, March 15). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, vol. 3, no. 160018, March 15, 2016, pp. 1-9. <https://www.nature.com/articles/sdata201618>.
25. Carroll, S. R. et al. (2020). The CARE Principles for Indigenous Data Governance. *Data Science Journal*, vol. 19, no.1. <https://doi.org/10.5334/dsj-2020-043>.
26. Global Indigenous Data Alliance (GIDA). (2023). *Indigenous Peoples' Rights in Data*. <https://www.gida-global.org/data-rights>.
27. See for example: Statistics Division. (2014). *Fundamental Principles of National Official Statistics*. United Nations Department of Economics and Social Affairs. <https://unstats.un.org/fpos/>.
28. See for example Mejias, U.A., & Couldry, N. (2024). *Data Grab: The New Colonialism of Big Tech and How to Fight Back*. The University of Chicago Press.
29. See for example Öhman, C. (2024). *The Afterlife of Data: What Happens to Your Information When You Die and Why You Should Care*. The University of Chicago Press.
30. European Union. (2016). Regulation (EU) 2016/679 of the European Parliament, Article 17. <https://gdpr-info.eu/art-17-gdpr/>.
31. Drawing upon the data life cycle as identified in GFC's earlier publication. World Economic Forum. (2023). *Data Equity: Foundational Concepts for Generative AI*, p. 6. <https://www.weforum.org/publications/data-equity-foundational-concepts-for-generative-ai/>.
32. United Nations. *Goal 13: Take urgent action to combat climate change and its impacts*. Sustainable Development Goals. <https://www.un.org/sustainabledevelopment/climate-change/>; Germanwatch. *Global Climate Risk Index*. <https://www.germanwatch.org/en/cr/>.
33. Data.org. (2023). *Innovative AI for Women's Financial Inclusion*. <https://data.org/stories/womens-world-banking/>.
34. Data.org. (2023). *Innovative AI for Women's Financial Inclusion*. <https://data.org/stories/womens-world-banking/>.
35. Two more examples focused on data equity in healthcare include mammography in detection of breast cancer in sub-Saharan Africa: Black, E. & Richmond, R. (2019). Improving early detection of breast cancer in sub-Saharan Africa: why mammography may not be the way forward. *Globalization and Health*, vol. 15, no. 3, 2019. <https://doi.org/10.1186/s12992-018-0446-6>; and identification of skin cancer lesions for people with darker skin: Lawson, A. (2024). Researchers teach AI skin-cancer diagnosis tool to see colour, *Brighter World*, March 4, 2024. <https://brighterworld.mcmaster.ca/articles/ai-skin-cancer-diagnosis-diversity/>.
36. Obermeyer, Z. et al. (2019, Oct 25). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, vol. 366, no. 6464, 2019, pp. 447-453. <https://www.science.org/doi/10.1126/science.aax2342>.
37. City of Boston. (2023). *Improving how the City of Boston collects gender data*. <https://www.boston.gov/equity-and-inclusion/improving-how-city-boston-collects-gender-data>; *Gender-aware guidelines and standards for City of Boston services*. <https://www.boston.gov/gender-aware-guidelines-and-standards-city-boston-services>.
38. Arellano Valdivia, J. (2024, February 22). *Data sovereignty, open mapping, and indigenous territories*. Humanitarian OpenStreetMap Team. <https://www.hotosm.org/updates/data-sovereignty-open-mapping-indigenous-territories/>. Indigenous Mapping Collective. (2024). <https://www.indigenousmaps.com/>.
39. Seamster, R. (2017, July 2). Creating maps that reflect indigenous geography. *Google Earth*. <https://blog.google/products/maps/creating-maps-reflect-indigenous-geography/>; Rush, T. (2017, June 21). Indigenous Lands in Canada are now in Google Maps. *Google Maps*. <https://blog.google/intl/en-ca/products/explore-get-answers/indigenous-lands-in-canada-are-now-in/>.
40. United Nations Economic and Social Council Resolution 2022/3 of 17 June 2022. "Ensuring that the work in the field of statistics and data is adaptive to the changing statistical and data ecosystem". [https://digitallibrary.un.org/record/3978014?ln=zh\\_CN&v=pdf](https://digitallibrary.un.org/record/3978014?ln=zh_CN&v=pdf).



---

COMMITTED TO  
IMPROVING THE STATE  
OF THE WORLD

---

The World Economic Forum, committed to improving the state of the world, is the International Organization for Public-Private Cooperation.

The Forum engages the foremost political, business and other leaders of society to shape global, regional and industry agendas.

---

**World Economic Forum**  
91–93 route de la Capite  
CH-1223 Cologny/Geneva  
Switzerland

Tel.: +41 (0) 22 869 1212  
Fax: +41 (0) 22 786 2744  
contact@weforum.org  
www.weforum.org