# DATA MANAGEMENT, GOVERNANCE: FOUNDATIONS FOR ACHIEVING THE GLOBAL BIODIVERSITY FRAMEWORK

# Supporting Environmental Policy and Governance

## Data Management, Governance: Foundations for Achieving the Global Biodiversity Framework[1]

Authors: Aditya Agrawal and Balakrishna Pisupati

2023

Contents



[1] This is a working document

## 1 Introduction and Background

Access to information as a right has gained credence in several countries around the world since the 1990s. **Data is an invaluable asset that is reusable, replenishable and infinite and has much importance in all aspects of governance**. In fact, international human rights bodies such as the UN Human Rights Committee, the European Court of Human Rights, the Inter-American Court of Human Rights and the European Committee on Social Rights have off late, accepted the existence of a right to information in support of equity and justice[1].

Recent instance of such recognition in international law is that of the Regional Agreement on Access to Information, Public Participation and Justice in Environmental Matters, also known as the '*Escazú Agreement*' adopted in March 2018 for Latin American and Caribbean countries, with the aim of promoting transparency and greater democracy in issues of environmental significance[2].

Access to substantial, good quality and reliable data pertaining to natural resources is more relevant in order to overcome the various environmental challenges facing the world. With the advent of electronic age and the current state within the 4$^{th}$ Industrial Revolution, collection of such data, its verification as well as curation and dissemination are not a difficult enterprise. In this context, both developed and developing countries must take advantage of being knowledge economies through their ability to collate and use the data they possess in various sectors, to support sustainable development.

However, at present, **there are significant gaps in understanding data management and governance issues from access, use and sharing perspectives**. This is a significant challenge to be addressed since using environment related data and information, with appropriate safeguards, would ensure local level development in a sustainable manner[3]. Key among current challenges is the lack of clarity and guidance regarding the protection of knowledge and information and related legal and regulatory frameworks.

Given the challenges in data management, there is, therefore, a need to undertake guidance to stakeholders on data management and governance issues. This is possible when stakeholders

understand data and information governance issues and standards. **While several principles and standards exist on data governance, their implementation at an institutional level remains weak, especially within the government and civil society sectors**. Equally important is the stakeholder engagement process and organization required to bring in the right expertise, experience, and perspective on making the right decisions on what data governance tools and approaches should be applied.

The ongoing digital revolution can be used to combat the triple planetary crises of dealing with climate change, nature protection and managing pollution. These are elucidated in the UN Secretary General's roadmap for digital cooperation[4]. The 8 key areas of action in this report can be achieved by elaborating and defining data management and governance principles and approaches. UNEP's role within the SG's strategy is to implement digital transformation within its work processes and to ensure that Member States have access to data and information that they need for implementing and reporting on environmentally sustainable development.

Thus, **there is a need to assess the current data and information governance systems, their modes of operations and provide a set of principles based on which future governance systems can emerge.**

The paper is divided into **two sections**, the first providing an overview of key issues and processes to be considered while working on data and information governance and the second a set of principles and standards to be considered by relevant stakeholders.

For purposes of immediate use of the processes, principles and standards elaborated in this paper, we have used the ongoing discussions under the Convention on Biological Diversity (CBD) through the recently adopted global biodiversity framework (GBF) as a starting point. **With 196 Parties to the CBD set to update and align their National Biodiversity Strategies and Action Plans (NBSAPs), the use of information provided in the paper can inform the way in which countries and stakeholders will generate, share, use and assess the data and information related to biodiversity.**

In addition, the paper is expected to inform work focusing on issues of digital transformation by institutions as well as for other national, bilateral and multilateral processes[5,6].

## 1.1 Global Biodiversity Framework

Countries from around the world gathered at the Convention on Biological Diversity (CBD) Conference of Parties (COP) 15 meeting in Montreal, Canada in December 2022 to agree on the Global Biodiversity Framework[7]. The Framework includes 23 targets to halt the loss of biodiversity by 2030, while achieving recovery and restoration of biodiversity by 2050. The Framework aims to facilitate implementation primarily through activities at the national level with support to subnational, regional, and global levels. In doing so, the underlying assumption is that a whole-of government and society approach is needed to achieve the goals over the next 10 years and set the foundation needed for the 2050 vision.

The GBF consists of 4 long-term goals for 2050, while each includes several milestones to assess progress by 2030. The 4 goals are:

a) The integrity, connectivity and resilience of all ecosystems are maintained, enhanced, or restored, substantially increasing the area of natural ecosystems by 2050; Human induced extinction of known threatened species is halted, and, by 2050, extinction rate and risk of all species are reduced tenfold and the abundance of native wild species is increased to healthy and resilient levels; The genetic diversity within populations of wild and domesticated species, is maintained, safeguarding their adaptive potential. Nature's contributions to people are valued, maintained or enhanced through conservation and sustainable use supporting the global development agenda for the benefit of all.
b) Biodiversity is sustainably used and managed and nature's contributions to people, including ecosystem functions and services, are valued, maintained and enhanced, with those currently in decline being restored, supporting the achievement of sustainable development for the benefit of present and future generations by 2050.
c) The monetary and non-monetary benefits from the utilization of genetic resources, and digital sequence information on genetic resources, and of traditional knowledge associated with genetic resources, as applicable, are shared fairly and equitably, including, as appropriate with indigenous peoples and local communities, and substantially increased by 2050, while ensuring traditional knowledge associated with genetic resources is appropriately protected, thereby contributing to the conservation

and sustainable use of biodiversity, in accordance with internationally agreed access and benefit-sharing instruments.

d) Adequate means of implementation, including financial resources, capacity-building, technical and scientific cooperation, and access to and transfer of technology to fully implement the Kunming-Montreal global biodiversity framework are secured and equitably accessible to all Parties, especially developing countries, in particular the least developed countries and small island developing States, as well as countries with economies in transition, progressively closing the biodiversity finance gap of 700 billion dollars per year, and aligning financial flows with the Kunming-Montreal Global Biodiversity Framework and the 2050 Vision for Biodiversity.

These goals are supported by 23 targets with time bound actions. While many of the goals and related targets are dependent on data to better understand, track and monitor progress, the word 'data' only appears once in the GBF.

### 1.2 National Biodiversity Strategies and Action Plans

As provided in Article 6 of the CBD, the National Biodiversity Strategies and Action Plans (NBSAPs) are the vehicles for countries to strategize and develop action plans for implementation of the Convention. With increasing focus on inclusive actions to achieve the objectives of the CBD, through the GBF there is a strong consensus among different stakeholder groups supporting the CBD to use the recent developments in data and information for decision and policy making at various levels.

There is no comprehensive global guidance on defining the data and information governance policies for accessing, using, sharing, and analyzing biodiversity data. Article 6 of the Convention on General Measures for Conservation and Sustainable Use states that each Contracting Party shall, in accordance with its conditions and capabilities:

(a) Develop national strategies, plans or programs for the conservation and sustainable use of biological diversity or adapt for this purpose existing strategies, plans or programs which shall reflect, inter alia, the measures set out in this Convention relevant to the Contracting Party concerned

(b) Integrate, as far as possible and as appropriate, the conservation and sustainable use of biological diversity into relevant sectoral or cross-sectoral plans, programs and policies.

Article 26 and Article 10(a) are closely linked to Article 6. The first calls for Parties to present, through their national reports, information on measures which have been taken for the implementation of the provisions of the Convention and their effectiveness in meeting the objectives of the Convention. The latter encourages Parties to integrate consideration of the conservation and sustainable use of biological resources into national decision-making.

NBSAPs generally include goals, targets, and actions as well as indicators for monitoring progress. These are supported by implementation plans as well. Therefore, it is important that the monitoring framework presented in the GBF be integrated into these NBSAPs in a manner that is relevant and enabling.

## 2 Digital Transformation

Through the advent of the internet and computers leading to the digital revolution, unprecedented rates of data acquisition, processing and dissemination was enabled completely transforming how we access and use data and information for everyday purposes. This has led to many positive changes including increased access to information, greater connectivity, and improved efficiency across many industries. Emerging technologies including the advancement of artificial intelligence (AI) and machine learning, Internet of Things (IoT), robotics, 3D printing, quantum computing, biotechnology and 5G and other advanced networking technologies are providing great opportunities in support of sustainable environmental management. However, technological interventions have also raised many concerns and questions about the impact of technology on societal values, privacy and security, and how digital technologies like social media are spreading misinformation.

While all the above have the potential to provide profound impact to society, environment and the economy, it does raise ethical and societal questions that must be governed in order to have transparency, understanding of expectations and possible outcomes, and offer the right tools and measures to protect people and planet. We must also acknowledge that there is a divide in

the digital economy between developed and developing countries, and that this offers its own challenges when it comes to the management, use and distribution of data, information, and knowledge.

Digitalization: **Use of digital technologies to turn products and services into a digital format to drive efficiency and innovation.**

Digital transformation: **Systems-level economic, societal and environmental transformations triggered as a result of digitalization where embedded technologies also drive efficiency, agility, culture and experience.**

## 3 Data Governance

**Data governance refers to a set of guidelines, policies, standards, and procedures that are used to manage and oversee the collection, storage, use and dissemination of data** (for purpose of this paper, we will focus on data related to biodiversity). The goal of applying data governance for biodiversity is to ensure data is collected, managed and used in a way that is accurate, consistent, reliable, secure and ethical. Data governance can improve both internal and external communication around the intended uses of data, increase value and reduce costs, and help an organization mitigate against any uncomfortable confrontations in the future from stakeholders on issues related to ethics, privacy, costs, bias, appropriate use, which will undoubtedly arise.

Components of a data governance framework can address the following:

1. **Policies and procedures**: rules and guidelines for how data should be collected, stored, used and distributed.
2. **Data standards**: these standards can define the formats, quality standards, metadata and sharing of data that must be followed when using data internally or externally to an organization.
3. **Roles and responsibilities**: Identification of key stakeholders and their roles and responsibilities in the data value chain and what this means for their role in data

governance. This may include data stewards, data owners, data custodians or data users.

4. **Process and procedures**: The process and procedures used to ensure the effective and efficient management of data within the organization.
5. **Tools and technology**: Tools and technology such as data quality tools, data governance platforms, data dictionaries, or APIs.
6. **Communications and training**: The mechanisms put in place to ensure all stakeholders are aware of and can build capacity on the data governance polices, process and standards for the organization.

**By having a data governance framework in place, a level of transparency is achieved that sets the rules for how data and information products are collected, managed, created, distributed, and used.** Countries, with good reason, will often question why external or global data should be used, and more importantly, why they should give their national data to a global institution. They may be reluctant to sign data sharing agreements. They will need reassurances that sharing data will not interfere or compromise their sovereign data, and that any data exchange occurs within a secure environment.

Data products and services will also have downstream implications. For example, a key issue with biodiversity data is around sensitive, threatened, or endangered species. By making location data on species occurrence available, there is the potential that these data can be misused. Therefore, data privacy and methods around generalizing location becomes very important for the continued protection of these species. Metadata, or information that describes the data, also becomes important so that the user understands how data was collected, methods applied, limitations of the data and how it can be applied.

### 3.1 Benefits of Data Governance

Several benefits[8] exist when a data governance framework is applied:

1. **Improved Experience and Services.** Through interoperability and connected systems, the public sector is enabled to provide a better experience to its citizens.

2. **Data-Driven Policy Creation**. Policy needs to be informed by data. When data is well-governed, it enables users to better understand the drivers and this means for public policy and how impact can be measured.
3. **Efficient Access to Data**. Users across sectors need access to data. Making data more accessible means that people have to submit data only once. This drives efficiency, an understanding of the true source, and therefore, saves time and money by enabling the automation of services.
4. **Enable Information Sharing**. Standardizing data formats leads to better collaboration and interoperability between systems. By having agency data in a single location, data duplication is eliminated and data quality can be further ensured.
5. **Create Transparency**. Engagement and collaboration with stakeholders across agencies, citizens and companies builds trust through transparency. Through data governance, public sector entities can publish data ethically through open data portals.
6. **Accountability and Integrity.** By creating transparency, you also want to create accountability on various datasets across the organization, especially in regards to the quality and what this means for uptake and usage of these data. As a result, processes can become more efficient leading to higher degree of integrity in the data.
7. **Modernize Data Systems**. Data governance can enable more structured and secure data. Through APIs and connected systems, it's easier to understand what data is stored, where it is stored, who has access, how it is being used and how many times it is used – thus providing metrics on the value of different datasets.
8. **Stay Current**. A data governance framework helps to identify potential problems and lays out a process for how these problems are communicated and resolved. It will also help identify outdated or deprecated datasets that are not fit for wider consumption.
9. **Reduce Fraud**. Fraud risk can be mitigated by aggregating data across registers to ensure consistency. Analytical tools can then detect when funds are going to an incorrect recipient thereby preventing fraud or mistakes in transfers from taking place.

### 3.2 Technical and Programmatic Considerations

At one level, data governance relies on the adoption and deployment of a set of principles, standards, schemas or technology. However, the successful implementation of a data governance approach also depends on leadership, culture, process, and collaboration.

Most governance programs today are ineffective because they depend on top-level leadership who often don't recognize importance and value creation of data governance[9]. As a result, data governance often becomes a set of policies or guidance implemented as a support function by IT and not often followed. Effective data governance requires rethinking organizational design and engaging with a range of stakeholders internally and externally to the organization.

### 3.3 Data Value Chain

Data management and governance is required across the entire data value chain. The data value chain refers to the sequence of steps that are involved in creating value from data. This can include a range of activities, from collecting and storing data, to cleaning and preparing the data for analysis, to using the data to make decisions, generate insights, or create new products or services. The data value chain can be thought of as a pipeline or process that transforms raw data into valuable information or knowledge.

The different stages of the data value chain can vary depending on the specific context and goals, but some common stages might include:

1. **Data collection**: This is the first stage of the data value chain, and involves gathering data from a variety of sources, such as sensors, databases, or surveys.

2. **Data storage**: Once data has been collected, it needs to be stored in a way that is safe, secure, and easily accessible. This can involve using a data warehouse, database, or other type of data storage system.

3. **Data cleaning**: Raw data often contains errors, inconsistencies, or other problems that can make it difficult to use. Data cleaning involves a range of processes for identifying and fixing these problems, such as filling in missing values, removing duplicates, or standardizing data formats.

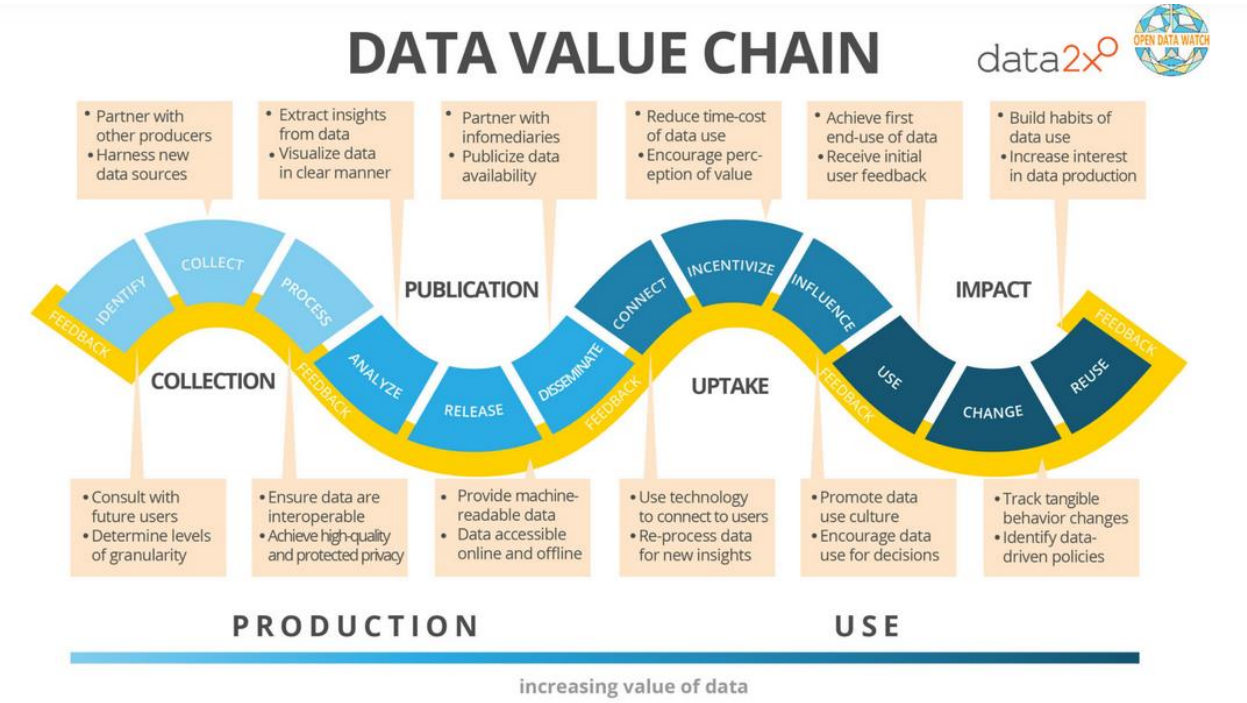
4. **Data analysis and use**: Once the data has been cleaned and prepared, it can be used for a variety of purposes, such as generating insights, making predictions, or identifying

patterns or trends. Data analysis often involves using specialized software tools or statistical methods to extract meaning from the data.

5. **Data dissemination**: Share the results of the analysis with others, either through reports, presentations, or other means. This can help to inform decision-making, generate new ideas, or support research and other activities.

6. **Uptake:** Once the data is released, you want to encourage the uptake of these data by connecting with users, incentivize through the added value of data to a particular problem, and have data influence culture and decision-making.

7. **Impact:** Finally, you want these data to have impact through usage by end-users, understand the change that is occurring due to the availability of these data and what this means for policy, and ensure data has multiple uses.




*A visual representation of the data value chain created by Open Data Watch for Data2X: https://opendatawatch.com/publications/the-data-value-chain-moving-from-production-to-impact/*

Overall, the data value chain is a series of steps that are involved in creating value from data. By understanding and managing the different stages of the data value chain, organizations can extract maximum value from their data assets.

### 3.4 National, Regional and Global Data Flows

Integration of the GBF into the NBSAPs at the national level will likely have the largest impact in terms of how the GBF is implemented and what this means for the protection and sustainability of biodiversity. However, at the national level, countries should also incorporate efforts at the local level to ensure biodiversity information being collected and managed by cities, academic and research institutions, private sector, and civil society organizations are also accounted for in the national system.

There is often a disconnect at this level due to often different data governance procedures in place at these levels and what is considered “authoritative” data or not – a definition often put into place by governments. As part of a multi-stakeholder effort, these local institutions should also be brought into the process for defining data governance and how biodiversity will be accounted for and managed across the country – something that should be made clear in the NBSAP.

Processes and standards should also be considered at the national level that allow for the easy transfer of these data to regional and global systems. What data, at what frequency and for what purpose is a process that will need to be determined by those regional and global institutions and the national entities. Global institutions like the Global Biodiversity Information Facility (GBIF) set clear standards, processes and procedures for how data from national entities can be incorporated into their global archives. Of course, the benefit here is that by opening up national data to a wider audience, the understanding of key issues, challenges and opportunities on biodiversity globally are better known and addressable.  It is also these more local and national data that really add value to broader global understanding.

# 4 Guidance on Developing a Data Governance Process

The following sections provide high-level guidance on the development of a data governance framework and process. It's important to note that this cannot be a 'one size fit all approach' as there will be unique circumstances across countries that require different approaches. To get the right governance framework established, it is much more about the process, the engagement that take place, and what this means for the organizational framework that is put into place that provides both strategic/programmatic and technical oversight.

## 4.1 Stakeholder Engagement

At the programmatic level, the goal is to establish the process of data governance through engagement with stakeholders across sectors. This cross-section of sectors is needed in order to better understand all the relevant parties that are producers or users of biodiversity information, and what this means for how these data are made available and used. While each entity may have different rules internally for how their data can be made accessible, it is important to open the dialogue as a way to build bridges across sectors, gain a better understanding of what initiatives and programs are underway, and how these parties can come together to advance biodiversity issues for the country.

Therefore, it is important to convene the relevant stakeholders through a systematic process that establishes the organizational structure for how these entities will cooperate, coordinate, and provide guidance. One approach is to develop a national workshop that is both multi-stakeholder and whole-of-government where relevant agencies and other institutions can discuss their work on biodiversity data, policy implications, latest innovations and key issues, challenges and needs. This is also a good opportunity to discuss the GBF more broadly and its implications for the development of an updated NBSAP.

By creating this inclusive process, it will open many perspectives across stakeholders in their understanding of biodiversity activities across the country. It will then also offer an opportunity to develop a governance body that includes both strategic and technical elements. The former provides overall strategic guidance and policy on how biodiversity data should be governed,

while the latter focuses on the technical implementation including the technology, standards, and other tools.

Therefore, **designing a governance framework that takes a top-down and bottom-up approach is suggested**. The top-down provides more of the strategic advisory function and builds broader buy-in and support for the program, while the bottom-up ensures alignment with existing programs and initiatives, maintains open communication avenues with users for needed feedback and provides technical expertise. Most importantly, it provides an inclusive platform where stakeholders are heard, part of the process, and provide their expertise for how data can be managed across the country.

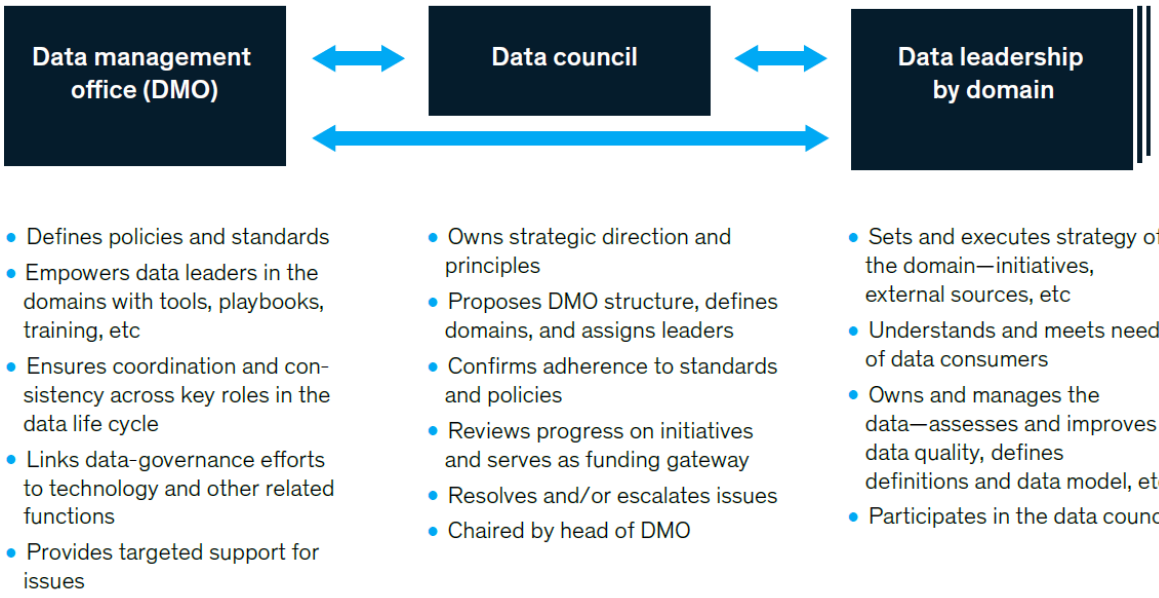
### 4.2 Organizational Structure

While there are many different approaches in how a governance process can be established, which largely depends on the culture, politics and institutional arrangements, there are some general approaches that can be considered:

1. Consider the creation of a central function for data management that includes the equivalent of a Chief Data Officer where the data strategy is developed and implemented. While a central function consolidates data management functions, the operational model can still be distributed across departments, agencies, and partner institutions.
2. Create a governance board that includes leaders and experts that can guide the overall direction on how implementation should occur and provide insights on priorities, engagement and issues that need to be considered internally and externally.
3. Define roles across stakeholder agencies organized by data domain such that users and producers are well defined and who should maintain custodianship around particular data types.

For example, a useful model developed by McKinsey and Company is provided in the following [7]:

A best-practice data-governance model typically includes three organizational components.




A similar, but alternative and higher level action could be to establish two main governance bodies:

**Data Governance Board:** includes high-level representation from government, private sector and civil society organizations. This can include national, regional, and international institutions and should be multi-stakeholder in nature. The main activities this group would engage in are the following:

- Provide strategic guidance
- Build political buy-in
- Identify partnerships
- Provide policy alignment
- Support advocacy and outreach
- Support fundraising

**Technical Advisory Group (TAG):** provides a more bottom-up function and multi-stakeholder with inclusion of national institutions, private sector, civil society organizations, domain experts and research institutions. This would be a much more technically oriented group with a focus

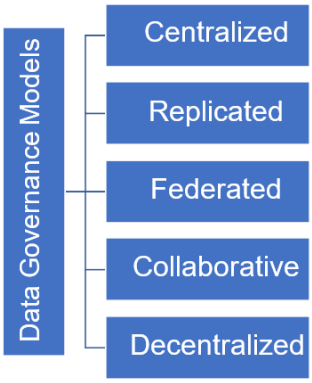
on entities working on biodiversity issues in the country, with regional and international support as needed. Key functions of the TAG would be to:

- Provide technical expertise
- Provide recommendations on technology, policy, principles and standards for effective data governance
- Provide guidance on work programme priorities
- Identify opportunities for alignment with existing program, projects, data and technology
- Create Communities of Practice and Working Groups (as needed)
- Identify emerging trends, issues and areas for innovation

**Depending on the context including the type and size of organization, capacity, mandate and users, the data governance model as described above can be articulated and shaped so that it is fit-for-purpose.** This requires engagement with internal and external stakeholders to create a model that has buy-in, is inclusive, and has the right level of transparency and accountability to ensure the governance is being responsive to the needs and priorities of the organization and its stakeholders.

### 4.3 Data Governance Models

There are different approaches to data governance, with some being more centralized than others. Individual organizations need to determine what will work best for them, keeping in mind the purpose for which data is being collected and used. For example, an NSO may want to develop a more centralized model for data collection, standard-setting, validation and security, given its role in coordinating the overall production and dissemination of official statistics at the national level. A more decentralized or more modular model of data governance may work better in instances where control over data is distributed. Too much decentralization does not work well in volatile environments that require data standards and coordination to tackle global information sharing challenges. Conversely, too much centralization can hinder experimentation and the creativity needed to innovate and to respond to emerging needs of data users and the quickly changing technological landscape.




A middle ground can be found in so called “replicated” and “federated” governance frameworks. The former is when a common data governance model is adopted (usually with only minor variations) by different organizations. The latter is when multiple organizations coordinate to maintain consistency across their data governance policies, standards and procedures, although with different schedules based on their level of engagement, maturity and resources.

A replicated data governance framework is well suited to promote interoperability across independent organizations and loosely coupled data communities, each of which has ownership over specific data assets. However, this kind of governance framework requires very clear institutional and technical mechanisms for communication and collaboration, including the provision of adequate incentives for the adoption of open standards and common data and metadata models, classifications, patterns for the design of user interfaces.

A federated governance framework allows multiple departments or organizations, none of which individually controls the all the data and technological infrastructure, to constitute a decentralized but coordinated network of interconnected “hubs”. Such “hubs” consolidate and provide a consistent view of all the data assets available across the network, reducing the complexity of data exchange management, and provides a space where disparate members of that network can engage with one another. Moreover, although the federated model provides a coordinated framework for data sharing and communication, it also allows for multiple representations of information based on the different needs and priorities of participating data communities. It leverages technology to enable collaboration and the implementation of common data governance mechanisms.

Collaborative approaches to data governance exist between organizations and institutions. and can be an effective way to engender a more multi-stakeholder, open and ecosystem approach to the tackling of interoperability problems. The benefits of collaborative approaches also include greater adaptability and flexibility than the more formal models mentioned above. The Collaborative on SDG Data Interoperability is one such example, as are the Health Data Collaborative3, the Committee on Data of the International Council for Science (CODATA) 4 and

even more formalized international standards organizations such as W3C5 and ISO6.

A level below that of governance frameworks sit business processes; series of tasks and activities that collectively result in the delivery of a product or service. Here, interoperability can play a role in helping gel the various parts of the business process together. The Generic Statistical Business Model (GSBPM) is a case in point.

### 4.4 Data Governance Roles

There are several roles that are often defined in regard to data governance. These roles include:

1. **Data governance board, committee or council:** This will involve a group of senior leaders and key stakeholders who are responsible for setting the overall direction and strategy for how data governance is developed and implemented across the enterprise.
2. **Data owner:** This is an individual or group that is responsible for the overall management and stewardship of a particular set of data within the organization. The ensure the protection of data through the implementation of policies, guidelines and MOUs to ensure quality, integrity, security and make decisions on its appropriate use and accessibility.
3. **Data steward:** The data steward is responsible for implementing the data governance policies and procedures within the organization. They implement data standards, monitor data quality, develop documentation about the data (metadata), enforce policy and handle any inquiries about the data.
4. **Data custodian:** This is an individual or group that is responsible for technically administering the data through management and operations. They are responsible for providing a secure infrastructure for maintaining the data, data archiving and security, backup and recovery, implementation of data access policies, and response time.
5. **Data user:** This is an individual or group that uses the data made available by the organization. They are responsible for following the rules, policies and procedures set in place regarding access and use of the data, and can report any concerns related to its use, management, or protection.

## 4.5 Data Access

At the local and national level, a key decision point that organizations will need to consider is what the data sharing model will look like for the data they hold. As many of the standards and principles will indicate in Appendix 6.1, the ideal is for data to be made open and free. This provides the most value regarding access and potential use of data with the understanding that data's value is dependent on its uptake and usage. By opening data, the downstream applications increase spurring further innovation, research, entrepreneurship and broader understanding and knowledge. From a societal perspective, making data freely available has benefitted both the individual and collective including not for profit entities, impacting theirs as well as the government entities' day-to-day decision making[10].

There have been many studies done over the years on the impact and value of open data. As one example, the US Geological Survey (USGS) opened Landsat satellite data in 2008. Prior to 2008, the cost for one scene ranged from $20-$4000 USD depending on the sensor. In this same period, less than 3000 images were ever sold in each month[11]. In 2009, the first full year that the data was made free and open, there were nearly 1 million downloads. In the years since, this has translated to 1.8 billion USD in benefits. Another study[12] estimated the economic benefit for the year 2011 as 2.1 billion USD across sixteen sectors. This is the potential power of open data in that it can dramatically increase the usage of data, which relates to both direct and indirect economic value based on the applications of the data.

However, there can be sensitive information, commercial interests or other reasons that make it difficult to fully open data. In these cases, for example, data sharing across government agencies and/or research institutions can be considered. Different business models can also be considered by offering a core where open data is made available for free, but value-added services and insights may come at a cost, suggesting hybrid approaches.

The arguments in either direction must be well thought out, and it is likely that those coming from different sectors will have different perspectives on the best approach. The reasons for not making data often come back to some kind of financial benefit, but as the example above provides, the overall impact of making data provides far more economic value than keeping it closed.

It is these types of discussions that **call for a data governance framework and organizational structure.** It offers a platform where these types of issues can be discussed across a range of stakeholders to determine the best path forward, and set the policies, rules and procedures accordingly.

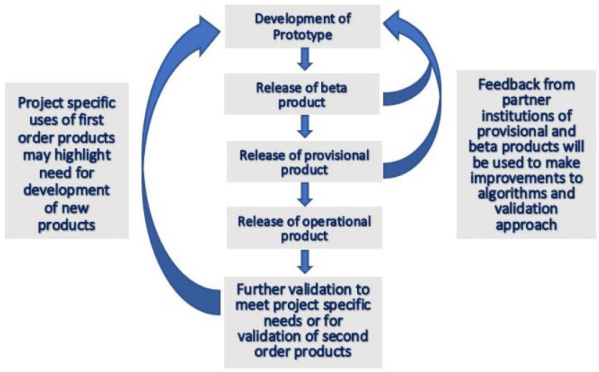
### 4.6 Technical Considerations

Data governance informs data management for which there are a number of technical considerations that must be considered and applied based on the context of the organization(s). These include the following:

1. **Data architecture:** The design for how data is managed (data modeling and design) – from collection through analysis, distribution, and consumption. The data governance framework should consider and align where possible to this architecture and fully consider all systems and data flows across the organization.
2. **Metadata:** Metadata is information about the data. It is critical from a user perspective to better understand how the data was produced, how it can be used, ownership and so on. Metadata is also important for data discovery through various platforms and search tools.
3. **Privacy and protection:** Privacy and protection of personal information and sensitive data must be accounted for in the data governance framework. This includes aspects related to data anonymization, data generalization and data protection.
4. **Systems integration and interoperability:** Related and inherent in the data architecture (as described above), how systems are integrated and what this means for data management, flow and usage across interoperable systems must be considered.
5. **Data quality:** A data governance framework should define and set processes for ensuring data quality and integrity. This can include the process for data cleaning and transformation, data validation or data standardization. Data quality information, including reference and master data, and measures must be transparent and regularly included in the metadata so data quality and accuracy is well understood by the user.
6. **Data security:** Data governance must include the various methods and protocols to protect sensitive data from unauthorized access.

7. **Storage and operations:** How data is stored and managed must be considered including physical storage via internal servers, or using cloud computing infrastructure. How data is backed up, what the recovery process looks like, data archiving and data deletion processes all need to be considered a part of the data governance framework.
8. **Accessibility:** How will data be made accessible internally and externally? Ensuring that data is "operational" or available routinely and reliably is important and whether these data will be shared internally, be made openly accessible, or something in between in regard to how users can get access all needs to be considered.
9. **Analytics and visualization:** How the data will be used should be considered. This included understanding how various systems across the organization as part of the broader data value chain are integrated and what this means for users and decision-makers across the organization to have the data in a form that allows for interpretation, reporting and visualization for its proper audience.

There are several standards, principles, and technical guidelines (Appendix 6.1) that can be applied to how biodiversity data is governed. Based on the context, stakeholders will need to make decisions on which standards, principles and processes apply best to their situation. The stakeholder engagement and organization structure should provide the process and pathway to make these data governance decisions.




The Digital Earth Africa program delivers operational earth observation products at scale for the entire continent. A data validation strategy was developed to move products from prototype to operational based on various validation methods and stakeholder engagement and feedback. By doing so, further trust and confidence was built in released products.

Again, the decisions that are made on how to deploy data governance at a technical level will very much depend on the local context. The framework should align to the overall data architecture of the organization. This includes understanding what data exists, the sources of these data, the various systems in place, dependencies across the organization and the data flows. Digital transformation, as described in Section 2, also is pertinent to understanding not only the technical architecture of an organization, but also its culture. In many cases,

organizations need to transform their way of doing things moving from legacy systems into more modern systems, something that the data governance approach should consider.

### 4.7 Coordination Across Global Institutions

Several global institutions were interviewed (Appendix 6.2) as part of this process to better understand existing initiatives and get expert guidance and feedback on data governance for biodiversity information. Each institution is advancing the use of biodiversity data including the use of standards, policies and principles that support data governance. However, there was a consensus that more needs to be done to better coordinate and collaborate across these institutions for how biodiversity data is managed and used.

While this is very common, it does highlight the need for a data governance approach at the global level that addresses data governance issues. By having such a mechanism in place, not only would it support how data is shared across these institutions, but most importantly, how data can flow between local, national, regional, and global systems.

For example, the Group on Earth Observations (GEO) is an intergovernmental partnership positioned within the World Meteorological Organization (WMO) that improves the availability, access and use of Earth observations. GEO promotes open, coordinated, and sustained data sharing and infrastructure for research, policy, and decision-making across a number of sectors. GEO includes over 100 member countries and 100 participating organizations (companies and organizations) thereby achieving a multi-stakeholder approach to its governance, while including a ministerial meeting every 4 years. GEO works with its membership to define a work program across flagship, initiative and pilot activities and includes a regional mechanism for engagement and uptake.

There are lessons learned from the GEO experience that can be applied to biodiversity data governance. In fact, GEO has a program focused on biodiversity observation networks (GEOBON, see Appendix 6.2). The intent would be to bring together global, regional, and national institutions to build better coordination and cooperation on the collection, management, use and distribution of these data.

## 5 Conclusions

This paper provides **a high-level approach with recommendations on possible standards, policies and principles that can be a starting point for the establishment of a data governance framework.** It is not meant to be prescriptive, but rather highlights the importance of data governance and provides an approach for how it can be established.

The intent is to further **refine the approach of this guidance by putting it into practice to better understand the issues, opportunities and constraints countries could experience**. Through these lessons learned, case studies will be developed outlining specific approaches countries have taken, which will only further inform this guidance document and global knowledge on addressing data governance for biodiversity. UNEP, with its institutional partners, looks forward to further engagement with country partners to develop data governance approaches more directly in response to the GBF, and create a model where this process can be replicated across multiple countries and regions.

**The development of a data governance framework for biodiversity data, information and knowledge is going to be essential to the successful implementation of the GBF and the development of NBSAPs in response**. Countries need to develop their own approach that is aligned to the capacity, culture, values, institutional arrangements, and technical architecture. It is important to create an inclusive and multi-stakeholder approach in how the organizational structure for governance is established and operated.

In support of this, UNEP need to consider development of guidance to countries on understanding, using and studying the impacts of data governance and management principles and approaches, primarily using the NBSAP revision/updating process. In addition, based on the experiences of working at country, regional and institutional levels, UNEP will focus on developing digital transformation governance policies, focusing on policy coherence, and supporting compliance to environmental decisions made through multilateral processes such as the MEAs and UNEA.

# 6 Appendices

## 6.1 Principles and Standards in use

### *6.1.1 Darwin Core*

Darwin Core is a standard maintained by the Darwin Core Maintenance Interest Group. It includes a glossary of terms (in other contexts these might be called properties, elements, fields, columns, attributes, or concepts) intended to facilitate the sharing of information about biological diversity by providing identifiers, labels, and definitions. Darwin Core is primarily based on taxa, their occurrence in nature as documented by observations, specimens, samples, and related information.

Reference and usage guides can be found at https://dwc.tdwg.org/.

From GBIF https://www.gbif.org/darwin-core:

The Darwin Core Standard (DwC) offers a stable, straightforward, and flexible framework for compiling biodiversity data from varied and variable sources. Originally developed by the Biodiversity Information Standards (TDWG) community, Darwin Core is 'an evolving community-developed biodiversity data standard. It plays a fundamental role in the sharing, use and reuse of open-access biodiversity data and today accounts for the vast majority of the hundreds of millions of species occurrence records available through GBIF.org.

In practice, using Darwin Core revolves around a standard file format, the Darwin Core Archive (DwC-A). This compact package (a ZIP file) contains interconnected text files and enables data publishers to share their data using common terminology. This standardization not only simplifies the process of publishing biodiversity datasets, it also makes it easy for users to discover, search, evaluate and compare datasets as they seek answers to today’s data-intensive research and policy questions.

Darwin Core is a widely used and accepted standard for biodiversity data and will be very applicable as part of the data governance guideline document for the GBF.

### 6.1.2 EML: Ecological Metadata Language

The Ecological Metadata Language (EML)[13] defines a comprehensive vocabulary and a readable XML markup syntax for documenting research data. It is in widespread use in the earth and environmental sciences, and increasingly in other research disciplines as well. EML is a community-maintained specification and evolves to meet the data documentation needs of researchers who want to openly document, preserve, and share data and outputs. EML includes modules for identifying and citing data packages, for describing the spatial, temporal, taxonomic, and thematic extent of data, for describing research methods and protocols, for describing the structure and content of data within sometimes complex packages of data, and for precisely annotating data with semantic vocabularies. EML includes metadata fields to fully detail data papers that are published in journals specializing in scientific data sharing and preservation.

EML is a widely used, community driven schema that supports documentation of data related to ecological research. It supports ecological and earth science data, and as discussed with the other institutions, there is a drive to have biodiversity and ecological data more integrated due to the obvious dependencies between each – something that the GBF does in fact incorporate.

EML is a useful model for standards around metadata such that researchers and users understand how data was collected and produced, as well as the underlying characteristics of these data. It is technical and largely geared towards the research community. Therefore, regarding its practical implementation at a national level and through governments responding to the GBF, it will be useful but will require much capacity development in its application.

### 6.1.3 Access to Biological Collection Data (ABCD) Schema

The Access to Biological Collections Data (ABCD)[14] Schema is an evolving comprehensive standard for the access to and exchange of data about specimens and observations (a.k.a. primary biodiversity data). The ABCD Schema attempts to be comprehensive and highly structured, supporting data from a wide variety of databases. It is compatible with several existing data standards. Parallel structures exist so that either (or both) atomized data and free text can be accommodated. Versions 1.2 and 2.06 are currently in use with the GBIF (Global

Biodiversity Information Facility, http://www.gbif.org/ ) and BioCASe (Biological Collection Access Service for Europe, http://www.biocase.org/ ) networks. Apart from the GBIF and BioCASe networks, the potential for the application of ABCD extends to internal networks, or in-house legacy data access (e.g., datasets from external sources that shall not be converted and integrated into an institution's own data, but be kept separately, though easily accessible). By defining relations between terms, ABCD is a step towards an ontology for biological collections.

As indicated previously, ABCD is in use by both GBIF and BioCASe. The standard is meant to be both comprehensive and general such that a broad array of concepts can be incorporated but the mandate is to only include the bare minimum set of elements to make the specification functional. ABCD is complex and highly structured but offers flexibility in its implementation and is a standard that can be considered by national institutions.

### *6.1.4 Creative Commons*

Creative Commons is a nonprofit organization that supports overcoming obstacles related to the sharing of data, knowledge, and creativity. They provide a set of licenses and public domain tools that allow for free, simple, and standardized ways to grant copyright permissions for creative and academic works, ensure proper attribution and allow others to copy, distribute and make use of the work. More information can be found at https://creativecommons.org/.

Creative Commons is a widely used, easily accessible and very straightforward approach for assigning copyright with freedom to share data. This license provides attribution and guidelines for sharing of data and is one that can be easily considered as part of the data governance framework.

### *6.1.5 FAIR*

The FAIR guiding principles for scientific data management and stewardship[15] include the following:

1. **F**indable: The first step in using data is to find them. Metadata and data should be easy to find for both humans and computers.

2. **A**ccessible: Once the data has been found, the user should know how to access these data.
3. **I**nteroperable: Data needs to be integrated with other data and interoperate with applications or workflows for analysis, storage and processing.
4. **R**eusable: The goal for FAIR is the reusability of data, and therefore, metadata and data should be well described in order to be replicated and combined.

Several principles further define each of these categories with examples and recommendations available for each. Accessible at https://www.go-fair.org/fair-principles/.

While FAIR has been gaining traction and is widely referred to and accepted among practitioners in the biodiversity informatics community, it is estimated that less than one percent of ecological data collected meet FAIR data principles[16]. In many cases, data are organized and stored from the perspective of the data collector rather than the downstream uses of these data across the community. This supply-driven approach is very common across academia, research institutions and intergovernmental organizations where understanding demand, use cases and user-centered design is not commonplace.

With that being said, we need to do more to gain the broader adoption and implementation of the FAIR data principles and integrating these, along with CARE, in the data governance for implementation against the GBF would make large gains. It would support policy development, and more importantly, the culture of sharing data across government and making data more open and accessible to wider range of users across private sector and civil society, which is critical.

### *6.1.6 CARE*

The CARE Principles for Indigenous Data Governance[17] are people and purpose-oriented reflecting on the role of data to advance indigenous innovation, culture, and values. They complement the existing FAIR principles and are put forth in similar format but focus more on the social aspects as opposed to the technical characteristics of open data and data sharing. The principles include:

1. **C**ollective Benefit: data ecosystems shall be designed and function in ways that enable Indigenous Peoples to derive benefit from the data.
2. **A**uthority to Control: Indigenous Peoples' rights and interests in Indigenous data must be recognized and their authority to control such data be empowered. Indigenous data governance enables Indigenous Peoples and governing bodies to determine how Indigenous Peoples, as well as Indigenous lands, territories, resources, knowledges and geographical indicators, are represented and identified within data.
3. **R**esponsibility: Those working with Indigenous data have a responsibility to share how those data are used to support Indigenous Peoples' self-determination and collective benefit. Accountability requires meaningful and openly available evidence of these efforts and the benefits accruing to Indigenous Peoples.
4. **E**thics: Indigenous Peoples' rights and wellbeing should be the primary concern at all stages of the data life cycle and across the data ecosystem.

CARE is very similar to FAIR but came into existence about 2 years after FAIR as a means to acknowledge indigenous culture and values. CARE principles should be seen as a required dimension of open and FAIR data[18] as described in the following diagram:

Operationalizing Complementary Principles: the Range of Be FAIR and CARE.

Regarding the GBF and its data governance framework, CARE applies to data that is indigenous, and where much biodiversity data is collected, it will likely need to be considered quite widely. Therefore, indigenous data needs to first be made FAIR and then CARE such that principles of CARE are well socialized across research communities and operationalized accordingly. The following diagram articulates the use of CARE across the data lifecycle [6]:

| Practice 'CARE' in data collection | Engage 'CARE' in data stewardship | Implement 'CARE' in data community | Use 'FAIR' with 'CARE' in data applications |
|---|---|---|---|
| Define cultural metadata<br><br>Record provenance in metadata | Use appropriate governance models<br><br>Make data 'FAIR' | Indigenous ethics inform access<br><br>Use tools for transparency, integrity and provenance | Fairness, Accountability, Transparency<br><br>Assess equity |

Implementation of the CARE Principles across the data lifecycle.

### *6.1.7 Open Data Charter*

The Open Data Charter (https://opendatacharter.net/) is a collaboration across 170 governments and organizations working towards open data based on a set of shared principles. The ODC promotes policies and practices that enable governments and civil society organizations to collect, share and use well-governed data. The Charter is adopted by 95 governments and endorsed by 78 organizations globally. The six principles of the ODC are as follows:

1. **Open By Default.** This represents a real shift in how government operates and how it interacts with citizens. At the moment we often have to ask officials for the specific information we want. Open by default turns this on its head and says that there should be a presumption of publication for all. Governments need to justify data that's kept closed, for example for security or data protection reasons. To make this work, citizens must also feel confident that open data will not compromise their right to privacy.
2. **Timely and Comprehensive.** Open data is only valuable if it's still relevant. Getting information published quickly and in a comprehensive way is central to its potential for success. As much as possible governments should provide data in its original, unmodified form.
3. **Accessible and Usable.** Ensuring that data is machine readable and easy to find will make data go further. Portals are one way of achieving this. But it's also important to think about the user experience of those accessing data, including the file formats that

information is provided. Data should be free of charge, under an open license, for example, those developed by Creative Commons.

4. **Comparable and Interoperable.** Data has a multiplier effect. The more quality datasets you have access to, and the easier it is for them to talk to each other, the more potential value you can get from them. Commonly agreed data standards play a crucial role in making this happen.
5. **For Improved Governance & Citizen Engagement.** Open data has the capacity to let citizens (and others in government) have a better idea of what officials and politicians are doing. This transparency can improve public services and help hold governments to account.
6. **For Inclusive Development and Innovation.** Finally, open data can help spur inclusive economic development. For example, greater access to data can make farming more efficient, or it can be used to tackle climate change. Finally, we often think of open data as just about improving government performance, but there's a whole universe out there of entrepreneurs making money off the back of open data.

### *6.1.8 The National Biodiversity Network – Data Exchange Principles*

The National Biodiversity Network (NBN, https://nbn.org.uk/) is a partnership founded on trust allowing for the exchange of wildlife information between individuals and organizations. The NBN Data Exchange Principles[19] were produced to create ground rules for taking part in the Network:

**Principle 1**

Biodiversity data should be easily accessible to enable their use for not-for-profit decision-making, education, research and other public-benefit purposes.

**Principle 2**

Making biodiversity data available should reduce the risk of damage to the environment. If it is likely to have the opposite effect, availability may need to be controlled.

**Principle 3**

Biodiversity data suppliers should make available sufficient meta-data to allow biodiversity data users to assess the scope and potential uses of their information holdings. When biodiversity data are supplied, accompanying information (meta-data) on its ownership, methods and scale of collection and limitations of interpretation, should be provided.

**Principle 4**

A clear transfer of authority should be made when a biodiversity data resource is put together, to allow biodiversity managers to act on behalf of the biodiversity data owners.

**Principle 5**

Managers of biodiversity data should make their framework of terms and conditions publicly available, allowing biodiversity data owners to have confidence that control will be exercised in the management and use of their data.

**Principle 6**

Personal data must be managed in accordance with the principles of the Data Protection Act 1998 and/or any subsequent legal provisions.

**Principle 7**

a) Managers and funders of biodiversity data should make basic facts freely available (except for handling charges if needed) for not-for-profit decision-making, education, research and other public-benefit purposes.

b) Biodiversity data suppliers should try to arrange resourcing of information provision so that charges for not-for-profit uses are minimal and charges for commercial uses are realistic but do not prevent the use of biodiversity data.

c) Biodiversity data users should expect to contribute to sustaining the provision of biodiversity data through contributing either in kind or financially to the collection, collation, and management of biodiversity data, or at the point of use.

The NBN data exchange principles provide very much integrated concepts of data governance and provide a real world example of execution against a national biodiversity network in the UK.

They serve as a practical tool to guide how participants participate and set some rules around that participation. This model and the case study of the NBN is one that can be further reviewed for applicability in other countries and serves as a good example for the guidance document on biodiversity data governance.

## 6.2 Institutional Initiatives

Several organizations and related initiatives that focus on biodiversity data were interviewed as part of this process. One thing that is clear from these interviews is that the landscape is fragmented in terms of coordination and collaboration across the biodiversity sector. The intent was to inquire further about existing initiatives, opportunities and challenges, and tools, resources and/or other references that should be considered in regard to biodiversity data governance.

### *6.2.1 Global Biodiversity Information Facility (GBIF)*

GBIF—the Global Biodiversity Information Facility (https://www.gbif.org/)—is an international network and data infrastructure funded by the world's governments and aimed at providing anyone, anywhere, open access to data about all types of life on Earth. As such, GBIF has done much work in terms of the standards and policies to make data open and sharable. Governments buy into GBIF to provide global benefits for making data as open as possible. In return, these data are often cited in research and policy tools giving credit and value to those contributing to these data.

GBIF also has a focus on better integration across agencies, capacity development and working on methods related to species distribution modeling. As such, GBIF is enriching its current data model to accommodate additional use cases including time series, ecological data and more components of biodiversity.

The key issue for GBIF, from a data governance perspective, is protection of sensitive species data. This includes providing guidelines on best practices for generalizing observation locations and the use of high versus low resolution data. GBIF is also incorporated the CARE principles to account for indigenous data sovereignty, culture and values.

In general, GBIF and its experience in open data, data sharing and providing tools and resources for countries and organizations to make their data available at a global level makes GBIF a strong partner in this initiative. Their guidelines and recommendations on standards that should be applied will be strongly considered as part of the guidance document to be developed.

### *6.2.2 GEOBON*

GEOBON (https://geobon.org), an initiative for developing Biodiversity Observation Networks under the Group on Earth Observations (GEO), has a mission to improve the acquisition, coordination and delivery of biodiversity observations and related services to users including decision makers and the scientific community. The vision is to have a global biodiversity observation network that contributes to the effective management policies for the world's biodiversity and ecosystem services.

GEOBON works with countries and institutions to establish national and regional biodiversity observation networks. They have developed a tool, BON in a Box, that is a regionally customizable and continually updated online toolkit for facilitating the start-up or enhancement of national or regional BONs. BON in a Box aims to serve as a technology transfer mechanism that allows countries access to the most advanced and effective monitoring protocols, tools and software thereby, lowering the threshold for a country to set up, enhance or harmonize a national biodiversity observing system.

GEOBON also works on the Essential Biodiversity Variables (EBVs), a process that is aligned to the GBF and provides direction for collection of data and transforming these data into indicators. Because GEOBON is part of GEO, the use of satellite and in situ data for biodiversity observations is a key aspect of this initiative, along with assessing biodiversity change over time – something well suited for satellite data.

GEOBON is working on its new strategic plan that will incorporate issues related to data privacy. Capacity development at the country level is a focus area for GEOBON understanding that there is a lack of capacity in producing data and moving these data into indicators.

GEOBON is doing much in the space of biodiversity indicator work, and because they are coming at this from the perspective of earth observation data, which makes applying various methods very scalable, GEOBON can offer valuable methods for collecting data and transitioning these into an indicator format. As previously noted, a key interest of GEOBON is to develop observation networks at the national level and using Bon in a Box could be a useful tool in regards to the GBF. Because GEOBON is part of GEO, there will be many references that can be used to support data governance.

### *6.2.3 NatureServe*

NatureServe (https://www.natureserve.org/) is a network organization based in the United States of over 60 organizations primarily in North America and 1,000 plus scientists serving authoritative biodiversity data. Its mission is to leverage the power of science, data and technology to guide biodiversity conservation and stewardship while envisioning a world in which the best available science informs conservation and stewardship decisions so that biodiversity thrives.

While primarily North America focused, NatureServe has a global footprint with a number of partners across Africa and other regions. Creating technology for data visualization, indicator dashboards and streamlined reporting services is a key focus area for the organization. It is supporting the CBD on analysis of indicators to track progress and has recently submitted a proposal in partnership with the Global Partnership for Sustainable Development Data (GPSDD) and Esri to create national level biodiversity indicator dashboards across 8 countries (6 in Africa and 2 in Latin America).

The focus for NatureServe is on issues related to data and technology for biodiversity. Similar to the above organizations, they have a wealth of experience when it comes to standards and building technology and have done much work on data collection and transforming data into indicator frameworks. While their work focuses primarily on North America, they have done considerable work at the global level, and if the funding is approved the indicator dashboard work identified above, this could be a good model for national implementation of the GBF.

In addition, NatureServe has much experience with data governance for biodiversity. As part of the discussion, they indicated that while there has been work done on data governance issues, it is complex, and therefore, fragmented. They agreed that there is not a unified framework for biodiversity data governance, that this is needed, but will be difficult. However, given all the various initiatives taking place in this regard, and the upcoming GBF, the timing is right to take this effort on.

### *6.2.4 UNEP-WCMC*

The UN Environment Programme World Conservation Monitoring Centre (UNEP-WCMD) is a global center of excellence on biodiversity and nature's contribution to society and the economy. UNEP-WCMC is a collaboration center between UNEP and the UK charity WCMC (https://www.unep-wcmc.org/en/about).

UNEP-WCMC has had a long history on the development of technology and bioinformatics tools in support of protecting the world's biodiversity. As such, they have developed several guideline documents in regard to the use of data and technology:

- Global Biodiversity Data Fitness Assessment – Progress toward establishing criteria for 'gold standard' decision-grade data[20]
- Biodiversity Platforms for Implementing a Sustainable World – Types and ingredients of effective biodiversity platforms[21]
- Scoping Planet+ Technical Architecture[22]

In addition to the above, UNEP-WCMC have noted the need Parties have expressed for national level capacity building and technical and scientific cooperation to enhance the generation, monitoring, reporting and assessment of data, information, and knowledge to support implementation of the GBF. As a result, UNEP-WCMC is consulting across a range of partners to determine the needs and opportunities for global knowledge support services for biodiversity in support of the GBF. The overall aim of this proposed initiative is to build on and interconnect existing tools, technologies, and networks in an inclusive manner, to support national efforts for

the implementation, monitoring, reporting and review of progress towards the agreed goals and targets of the post-2020 global biodiversity framework.

### *6.2.5 Biological Collection Access Service*

The Biological Collection Access Service (BioCASe, https://www.biocase.org/) is a transnational network of primary biodiversity repositories. It links together specimen data from natural history collections, botanical/zoological gardens, and research institutions worldwide with information from huge observation databases. The aim is to make the world's data on biodiversity data freely and universally accessible on the Internet through data portals and web services, a goal that BioCASe shares with related initiatives such as the Global Biodiversity Information Facility (GBIF) and Integrated Digitized Biocollections (iDigBio). In the past years, BioCASe has developed into a widely accepted standard for data sharing and has laid the foundations for several thematically - both taxonomically and geographically - specialized networks (see box to the right).

While BioCASe provides useful standards and tools for data sharing, it seems that these tools are more applied to large collections associated with academia, museums and research institutions collecting biodiversity data. The software and related tools require download and installation. Therefore, from a data governance perspective related to the GBF for implementation at a national level, it is less clear on how these tools could be applied. They offer good reference and good practice for data sharing, but likely less applied when considering how governments can utilize these tools.

## 6.3 References


[1] Maeve McDonagh, The Right to Information in International Human Rights Law, 13:1, Human Rights Law Review. 25-55 (2013).

[2] Alicia Bárcena, 'The Escazú Agreement: An Environmental Milestone for Latin America and the Caribbean, Economic Commission for Latin American and the Caribbean', (October 12, 2018, 12:13 PM) https://www.cepal.org/en/articles/2018-escazu-agreement-environmental-milestone-latin-america-and-caribbean

[3] Shyama Kuriakose, Balakrishna Pisupati Knowledge commons: Current issues related to natural resources and biodiversity. FLEDGE and FLS. 2019.

[4] Roadmap for Digital Cooperation: Secretary General's Report. 2020. https://www.un.org/en/content/digital-cooperation-roadmap/assets/pdf/Roadmap_for_Digital_Cooperation_EN.pdf

[5] Improving biodiversity: How can digitization help. https://www.epc.eu/content/PDF/2020/Digitalisation_v3.pdf

[6] Standards and Digital Transformation. https://www.unido.org/sites/default/files/files/2021-11/Standards%20and%20Digital%20Transformation_Complete_2021.pdf

[7] CBD/COP/15/L25, Kunming-Montreal Global Biodiversity Framework, 7-19 December 2022, Montreal, Canada, https://www.cbd.int/doc/c/e6d3/cd1d/daf663719a03902a9b116c34/cop-15-l-25-en.pdf

[8] Martha Bouchot, The Public Sector Needs Data Governance, *The Data Administration Newsletter,* 21 December 2022, https://tdan.com/the-public-sector-needs-data-governance/30228.

9 Bryan Petzold, Matthias Roggendorf, Kayvaun Rowshankish, and Christoph Sporleder, June 2020, Designing Data Governance that Delivers Value, McKinsey Technology.

10 Shyama Kuriakose and Balakrishna Pisupati. 2019 Knowledge Commons: Current Issues related to Natural Resources and Biodiversity. FLEDGE and FES, India.

[11] Zhe Zhu, Michael A. Wulder, David P. Roy, Curtis E. Woodcock, Matthew C. Hansen, Volker C. Radeloff, Sean P. Healey, Crystal Schaaf, Patrick Hostert, Peter Strobl, Jean-Francois Pekel, Leo Lymburner, Nima Pahlevan, Ted A. Scambos, Benefits of the free and open Landsat data policy, *Remote Sensing of Environment*, Volume 224, 2019, Pages 382-385, ISSN 0034-4257, https://doi.org/10.1016/j.rse.2019.02.016. (https://www.sciencedirect.com/science/article/pii/S0034425719300719)

[12] National Geospatial Advisory Committee Landsat Advisory Group, The Value Proposition for Landsat Applications, 2014, pgs 1-11.

[13] Matthew B. Jones, Margaret O'Brien, Bryce Mecum, Carl Boettiger, Mark Schildhauer, Mitchell Maier, Timothy Whiteaker, Stevan Earl, Steven Chong. 2019. Ecological Metadata Language version 2.2.0. KNB Data Repository. doi:10.5063/F11834T2 https://eml.ecoinformatics.org/

[14] Access to Biological Collections Data Task Group. 2005. Access to Biological Collection Data (ABCD). Biodiversity Information Standards (TDWG) http://www.tdwg.org/standards/115

[15] Wilkinson, M. D. et al (2016). The FAIR Guiding Principles for scientific data management and stewardship. Sci. Data 3:160018 doi: 10.1038/sdata.2016.18. Accessible via https://www.nature.com/articles/sdata201618

[16] Hackett, R. A., M. W. Belitz, E. E. Gilbert, and A. K. Monfils. 2019. A data management workflow of biodiversity data from the field to data users. Applications in Plant Sciences 7(12): e11310.

[17] Research Data Alliance International Indigenous Data Sovereignty Interest Group. (September 2019). "CARE Principles for Indigenous Data Governance." The Global Indigenous Data Alliance. https://www.gida-global.org/care

[18] Carroll, S.R., Herczog, E., Hudson, M. *et al.* Operationalizing the CARE and FAIR Principles for Indigenous data futures. *Sci Data* **8**, 108 (2021). https://doi.org/10.1038 https://www.nature.com/articles/s41597-021-00892-0

[19] The National Biodiversity Network DATA EXCHANGE PRINCIPLES. https://nbn.org.uk/the-national-biodiversity-network/archive-information/data-exchange-principles/

[20] Lauren V. Weatherdon, Bex Gottlieb, Ben Tregenna, Naomi Kingston, Neil Burgess, Corli Pretorius. 2021. Global Biodiversity Data Fitness Assessment: Progress toward establishing criteria for 'gold standard' decision-grade data. Cambridge, UK: UNEP-WCMC. 6 pp.

[21] Lauren V. Weatherdon, Jennifer Preston, Nina Bhola, Ben Tregenna, Naomi Kingston, Neil Burgess. 2021. Biodiversity platforms for implementing a sustainable world for nature and people: types and ingredients of effective biodiversity platforms. Cambridge, UK: UNEP-WCMC. 17 pp.

[22] Ben Tregenna, Lauren Weatherdon, Osgur McDermott-Long. 2021. Scoping Planet+ technical architecture. Cambridge, UK: UNEP-WCMC. 9 pp.