



**Taskforce on Nature-related
Financial Disclosures**

Nature-related Data Catalyst Initiative

September 2022 workshops –

Summary of learnings and next steps

framework.tnfd.global

Executive Summary

Data coverage – advancing comprehensive nature-data solutions is in some ways limited by the availability of data for particular realms (e.g. oceans), for remote locations (e.g. locations away from where organisations have their operations), and data across temporal scales.

Action – further collaboration and investment to ensure data are capable of meaningfully capturing the relevant variable/s, both static and dynamic. Standards will therefore need to be clearly defined to enable the collection of data that are relevant and useful, and at required frequencies for businesses and financial institutions.

Data standardization – Data are often not attributed, collated or presented in a standardised way which makes it challenging to ensure consistency in measurement between spatial locations or across a time series. Often, data are collected for very specific research questions that are not always relevant for the purposes it is being used for.

Action – development of a data standard will increase trust in datasets being collected, collated and compared. A standardised set of principles may also channel additional funding into the collection of additional data where gaps are identified. It would also improve the auditability of data.

Transparency – In order for data or models to be used by TNFD users or other data providers, the underlying data must be trusted, and its limitations understood. There is a need for more transparency on models used and limitations of assessment and data.

Action – increase transparency over data gaps, e.g. self-assessing datasets and highlighting where there is spatial/temporal uncertainty.

Data accessibility – Data accessibility was often mentioned by DC participants as another major limiting factor for the advancement of nature-related data, as there are often strict licensing restrictions in place, particularly for commercial use.

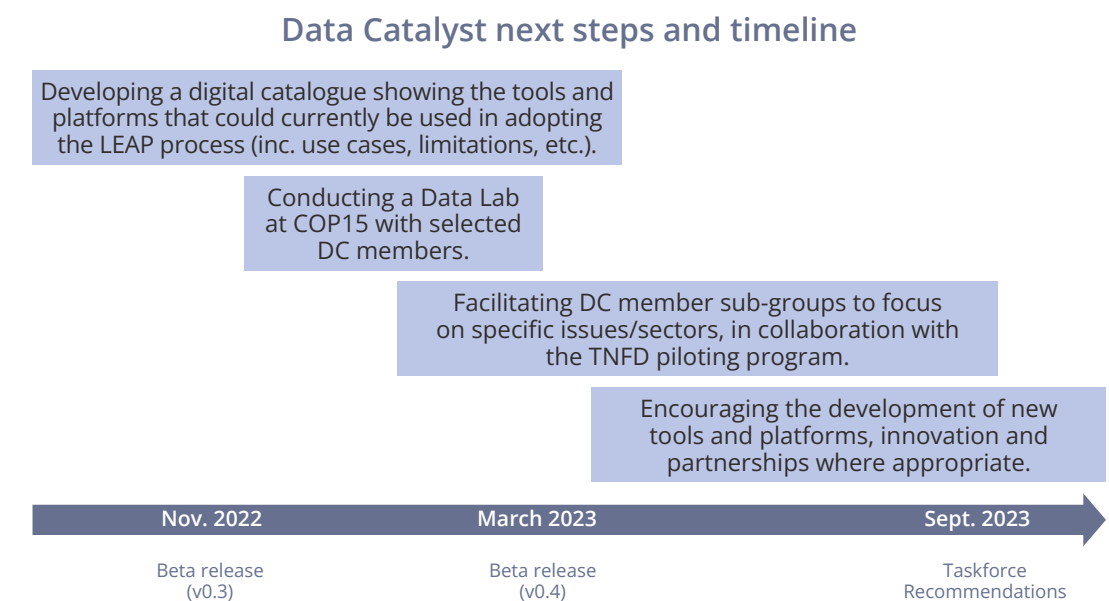
Action – increase collaboration between private companies, governments and NGOs to make data publicly available.

Relevance to decision making – It is important for the data providers and users to understand TNFD definitions of impact, dependency, risk and opportunity to be able to assess what is required and relevant, and to be able to do the analysis and decision-making in the lens of TNFD. There is also a need for more context, understanding and definition of what is nature positive, like we have for Net Zero. This would enable corporations and financial institutions to make a commitment, have a point of benchmark, establish baseline and measure progress. This would help clarify what types of data are relevant.

Action – distinguish between data required for different risks and opportunities. Guidance on how to conduct risk and opportunity assessment with regards to nature. Some pre-prioritisation and guidance developed by scientists and technical experts by sector, country, region, ecosystem, etc. of top impacts for example would be helpful to do the analysis. Transparency around how metrics and risks/opportunities were developed and identified will facilitate decision-making.

Next steps

Below is an overview of key next steps and timeline. See the last page of the document for more details.



Summary of learnings

A first round of the Nature-related Data Catalyst workshops was held the week of September 12th-16th 2022. Participants were split into groups of approximately 15 entities per session, allowing for these workshops to be representative of the nature-related data landscape, while ensuring proactive engagement from all the participants. Pre-workshop questions on data gaps were shared with participants in advance and each session lasted for 2 hours. The workshops were held under the Chatham House Rule, so all insights have been anonymised and aggregated where appropriate for onward dissemination. This document summarises key themes identified and discussed during the workshop, presents answers to the Q&A survey shared with participants in advance of the workshops and lays out the next steps for the participants.

The major themes uncovered through the Data Catalyst (DC) discussions validated the conclusions of the [TNFD Landscape Assessment of Nature-related Data and Analytics Availability](#) paper and included limitations brought on by inconsistent data coverage, the benefits of standardisation and transparency, the issue of data privacy, and how addressing each of these could impact the use of nature-related data in corporate decision-making.

Data Coverage

Data coverage was routinely selected by DCI participants as a major limiting factor affecting the utility of nature-related data for addressing the requirements of TNFD. Participants confirmed that advancing comprehensive nature-data solutions is in some ways limited by the availability of data for particular realms (e.g. oceans), remote locations (e.g. locations away from where organisations have their operations), and data across temporal scales. Below are a few insights mentioned by participants on these factors.

- Where proxies and modelling are used to fill in data coverage gaps, ground source data (e.g. site-level monitoring and/or data collection) is often, where practical, still required to validate model assumptions. Mobilising field surveys layered with remote sensing location-specific data will enable more accurate modelling but requires significant resource, logistical coordination and sensitive interaction with Indigenous Peoples and Local Communities (IPLCs) and rights-holders (e.g., where providing knowledge of species present and/or indicator species for a certain landscape as a measure of health).
- This will require further collaboration and investment to ensure data are capable of meaningfully capturing the relevant variable/s, both static and dynamic. Standards will therefore need to be clearly defined to enable the collection of data that are relevant and useful, and at required frequencies for businesses and financial institutions.

- Periodic efforts to catalogue and consolidate nature-related data across a broad coverage area (like what IPCC produces with the AR/CMIP) can demonstrate and highlight coverage gaps and guide research capacity towards addressing them. In some cases, data collectors cannot physically access the desired data collection sites. There are a number of reasons for this such as remoteness of location, ownership restrictions, or areas of conflict. This could be mitigated through the use of remote sensing combined with ground-based data collected by those with access to the sites who are trained and incentivised on how to collect the necessary data.
- New and existing technologies can and are being used to passively accumulate data (e.g., earth observation (EO) satellites; IOT devices/smartphones; sensors onboard commercial transportation). The continued advancement of these sensors in terms of capability (resolution/spectral coverage), accessibility and coverage (spatial and temporal frequency) combined with novel analytical methods to detect nature indicators (e.g., employing artificial intelligence and/or machine learning) will improve data availability, as systems are trained to enable detection ecosystem indicators. However, the challenge will be access to these data as the companies producing these data products will need revenue and therefore must charge for its use. This could limit access to only those organisations that have the resources to pay for it and/or prioritise the acquisition and development of data to support their assessments. It is also important to note that often the ancillary data that are used to model or classify EO data are often available only for non-commercial uses, so entities that are trying to derive a commercial data product may be restricted from using these ancillary data sets.
- Incomplete pictures of resources/ecosystem impacts stem from acknowledged data coverage issues. Understanding cumulative impacts is dependent on having a complete picture of ecosystem impacts and the various activities that are contributing to those impacts. For example, with water availability, if a reference condition for water volumes was defined that represented what would be considered 'normal' or 'sufficient' given current and future climatic conditions, then the cumulative effects on water availability could be assessed by considering water usage by all water users in a watershed. Whilst companies can and do measure their water consumption, the reference condition or baseline data with which to assess their cumulative impacts rarely exists.
- There is a need to develop partnerships and collaborations with different data providers and reinvest revenue back into the biodiversity datasets so they can stay updated but also remain or become available to a commercial audience. This will reduce temporal biases and other coverage issues. A good example of an effective cost recovery mechanism is IUCN's IBAT which directly supports the update and maintenance of three of the world's most authoritative global datasets: the World Database on Protected Areas, the World Database of Key Biodiversity Areas, and the IUCN Red List of Threatened Species.

Data Standardisation

A key barrier to using available data and data scalability relates to data standardisation. Participants recognised the benefits of ever-evolving data collection and analytical methods but raised concerns about how these changing approaches affect our ability to use historic data and create data solutions at scale. Data are often not attributed, collated or presented in a standardised way, which makes it challenging to ensure consistency in measurement between spatial locations or across a time series. Often, data are collected for very specific research questions that are not always relevant for the purposes it is being used for. DCI participants emphasised that standardisation is critical for baselining, ensuring scalability and enabling modelling as further explained below:

Baselining

- Standardisation of nature-related data over time will allow for an organisation to monitor change in state over time, which is crucial.
- This baseline is complicated due to constantly changing and improving methodologies.
- There must be a balance of old and new methods, and techniques for integrating old with new data sets, to ensure the value and usability of legacy data in baselines.

Scalability and modelling

- Empirical data are needed inform models and algorithms, but these data must be in a standard format in order to allow data to be aggregated and processed in a more scalable manner. Variance in approaches and methodologies for how empirical data are measured may lead organisations to arrive at a “cloud of plausibility” in their assessments, instead of arriving at a clear and precise answer. This may be a reality organisations must contend with if methodologies are not standardised.
- Some organisations are working to develop tools which aim to standardise data from distinct sets. A cloud-based geospatial analysis platform that enables users to visualise and analyse satellite images of Earth was provided as an example that has attempted to standardise diverse datasets. However, they have found challenges in that there is no consistent way to annotate datasets. Another DCI member provided an example of how GEDI’s LiDAR data has been utilised to derive biomass coverages from satellite imagery.
- Methodologies will evolve over time and users will converge on the methodologies that best suit them (e.g. produce reliable results that are variable, and use available inputs/accessible data). This convergence of methodologies will be accelerated when and if standards assessment and disclosure metrics are broadly adopted by the industry and financial institutions.

- Some Data Catalyst members are taking it upon themselves to create datasets which meet high standards or “decision-grade criteria” and can be regularly maintained, specifically for commercial use.
- Development of standards for how data providence is maintained will increase trust and allow verification of datasets. A standardised set of principles, and/or novel solutions such as tokenisation, for attribution of ownership and licensing nature-related data may also channel additional funding into the collection of additional data where gaps are identified.

Transparency

In order for data or models to be used by TNFD users or other data providers, the underlying data must be trusted, and its limitations understood. There is a need for more transparency on models used and limitations of assessment and data. Below are a few insights mentioned by participants on the importance of transparency.

- It is important to increase transparency over data gaps, e.g. self-assessing datasets and highlighting where there is spatial/temporal uncertainty. It was suggested that TNFD or similar organisation could provide a platform whereby organisations could upload datasets and connote gaps/limitations in data.
- Tagging data (metadata) is essential to allow sufficient data transparency to build trust and promote reasonable insights e.g. time stamps and geotagging. Metadata must include additional operations that have been applied to the data. A secure approach and controls for tracking ownership and licensing is required to ensure correct usage and revenue attribution. Without this, institutions may be hesitant to share the data or make it available for commercial uses. This data lineage will be of particular importance in commercial models and in order to allow monetary benefit to flow back to the original data collector.
- This will also allow greater use of citizen science data e.g. from camera traps that will reduce costs around data collection, although there will need to be more support to local communities on best practice for data collection to ensure this data is useful for decision-making.
- Timestamp data taken together with occurrence records would greatly improve predictive models. For example, the combined use of high-resolution remote sensing data and ground-truth data from Global Biodiversity Information Facility (GBIF) / Ocean Biodiversity Information System (OBIS) would improve the temporal resolution of prediction and may enable near-real time evaluation.
- Transparency around how metrics and risks/opportunities were developed and identified will facilitate decision-making.
- Issuer asset location including supply chain transparency should play an integral part in our priorities of data development over the coming months.

Data Accessibility

Data accessibility was frequently mentioned by DC participants as another major limiting factor for the advancement of nature-related data, as there are often strict licensing restrictions in place, particularly for commercial use. This issue is further discussed below.

- Open-source datasets are often created by institutions and organisations, but they often lack the funding or capacity to maintain these datasets. On the other hand, where datasets are treated as an asset which owners seek to capitalise, revenues are available to continuously maintain and improve the data products. Open data licensing will often preclude commercial uses, thereby limiting access to important data sets by commercial data providers.
- There is a lack of agreement in the data community over which operating model works best. Some advocate proprietary data as it fuels innovation, competition and options for end-users. There remains a question over what combination of public and private finance should pay for data collection, maintenance, and verification.
- In some instances, the data market is being saturated by duplicative primary datasets as companies are “re-inventing” the wheel.
- Some data providers are concerned that publishing their data will expose them to liability risk. This risk would be less acute if there were multiple open-source data sources.
- There is clearly a need to increase collaboration between private companies, governments and NGOs to improve data availability.

Centralising nature-related data platforms and quality

The pros and cons of a centralised nature-related data platform were discussed and summarised below.

- The development of a platform or ecosystem which brings distinct databases together to become the ‘one stop shop’ for biodiversity/nature data could reduce duplication of datasets, increase collaboration and dialogue between data providers. Indeed, a nature-related data access portal maintained by one organisation would allow a single place to go for data, making it easier for those seeking data and potentially increasing consistency. However, this approach would require extensive resources for its development and subsequent constant maintenance. There would also be challenges in managing relationships with many different data providers and in harmonising data from multiple sources.
- Therefore, a compromise proposal would be for TNFD (or other organisation) to develop and maintain a register of data sources (a digital catalogue) so that data/platforms can be easily identified, instructions for appropriate use, and where and how data can be accessed from data provider. In the catalogue, tools could be tagged against some criteria (sectors, geography, LEAP elements, etc.). In terms of evaluating the suitability of data tools and platforms, and the quality of available data, a

governance function should exist to curate the catalogue, by providing a definition for what constitutes suitable data for a given purpose, and evaluating data sets against these criteria.

Relevance to decision-making

In addition to addressing each of the themes previously discussed, a few other elements were discussed in order to facilitate the use of nature-related data in corporate decision-making.

- Aligning data products with TNFD will serve a dual purpose of enabling market uptake of the TNFD framework, which will in turn create further market demand for relevant nature-related data. It is therefore important for data providers and users to understand the TNFD assessment approach and definitions (such as impact, dependency, risk and opportunity) and how their data products can be used to support analysis and decision-making through the lens of TNFD.
- There is also a need for more context, understanding and definition of what is meant by Nature Positive. This would enable corporations and financial institutions to make a commitment, establish points of reference as benchmarks, establish baseline and measure progress. This would also help clarify what types of data are relevant. Similar to climate scope 1, 2 and 3, there is a need to make a distinction between what is within a company's direct control (impact, risk and dependencies related and location specific) versus what is outside of that control (opportunities in a broader landscape). Each of these elements require different data (location-based versus regional/national scale) so this would help clarify what types of data are relevant. (NB the concept of ‘scopes’ will be included in future TNFD releases).
- To avoid “perfect being the enemy of the good”, it is important to start now and take an iterative approach, working to improve upon and increase the utility of available data sources, while identifying and developing new data sources. This is especially important when it comes to translating impacts and dependencies into risks and opportunities, a key requirement in TNFD-aligned reporting. Given that different businesses and sectors will have a unique context, impacts and dependencies, there will need to be different ways and approaches to translate data into risks and opportunities. There will not be one single translation tool.
- Existing tools and processes that have been developed by NGOs and governments to assess and translate impact and dependency into risk and opportunity for nature and society, whereas corporations are most often treated as a source of impact to nature. Some of these frameworks and processes could be adapted and leveraged to the TNFD context, by considering both how corporations rely on natural assets and ecosystem services and how business activities can create opportunities that benefit nature and reduce impacts.
- It would be useful for TNFD to distinguish between data required for different risks and opportunities. Corporations will need guidance on how to conduct risk and opportunity assessment with regards to nature (such as the ENCORE tool). High-

level prescriptiveness from TNFD would also help, with some pre-prioritisation and guidance developed by academics and technical experts by sector, country, region, ecosystem, etc. defining the top impacts. Similar to the SASB materiality guidance by sector. This would help the risk and opportunity analysis and decision-making as well.

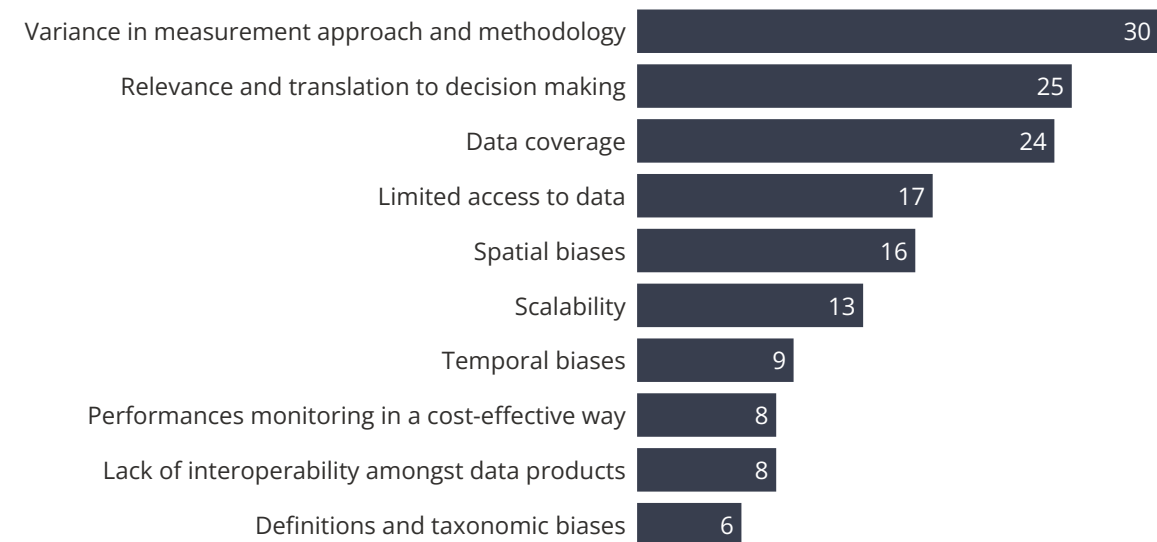
Below is a summary table including some of the above challenges and highlighting opportunities and recommendations by type of entity:

	Generation & Collection (contribution)	Storage, Access & Curation (contribution)	Analysis (contribution/consumption)	Communication
Entities	Biodiversity data contributors Asset data contributors Corporates	NGOs Governments Trade organisations Academia Non-profit organisations	Commercial data providers Financial Institution data consumers	Non-profit organisations Governments Inter-governmental organisations Trade organisations Data contributors
Challenges	Biased/inaccurate data collection resulting in data gaps for some taxa or region/s	Lack of access to some datasets due to licensing and privacy restrictions	Inappropriate analysis based on opaque superficial methodologies	Miscommunication by data contributors and/or misunderstanding by the data consumers
	Expensive, time consuming, labour intensive	Prohibitive cost of accessing the physical spaces housing biodiversity data and associated specimens	Scarcity of expertise in big data analysis with crossover expertise in biodiversity	Quantitative assessment of the value of biodiversity
	Incompatible data formats & platform	Financial cost of storage – some data behind paywalls	Incompatibilities between platforms & tools	Biased view of the proponents of biodiversity science
			Spatial/temporal uncertainty	Adverse political pressure
			Lack of access to some datasets	
			Relevance to decision making	
	Data often collected for very specific research questions		Prioritisation will vary e.g. it is location dependent so difficult to aggregate scores	Analysis must be effectively communicated to stakeholders and policy makers
	Data are often not attributed, collated, or presented in a standardised way			

	Generation & Collection (contribution)	Storage, Access & Curation (contribution)	Analysis (contribution/consumption)	Communication
Opportunities	Interdisciplinary opportunities	Innovative solutions through integration of different datasets	Development of intelligent algorithms for big data analysis	Test effective communication strategies
	Development of infrastructures for data aggregation	One-stop shop to access, integrate and analyse multiple datasets	Rapid decision based on robust data analysis	Development of professionals to straddle the science-policy interface
	Automated collection of high-volume data in some domains	Reusable data	Development of system-level solutions	Understand consumer needs and challenges via contributor/consumer collaborations e.g., case studies
	Data collection could be informed by analysis of previously available datasets by identifying specific data gaps	Research leveraging through multi-disciplinary collaboration	Multidisciplinary collaboration	
	Ground truthing importance i.e. not only relying on satellites	Improve transparency Infrastructural developments will drive innovative data collection Public projects acting as test runs where organisations can collaborate to innovate		
Drivers of change	Research leveraging through multi-disciplinary collaboration	Data monetisation	Lower data transaction costs	Outreach
	Increased prioritisation and demand	Reputation	Commercial opportunity	
	Transparency over data gaps	Consumer demand – regulation, frameworks, reputational risk	Consumer demand – regulation, frameworks, reputational risk Regulation	
Recommendations	Development of data collection protocols to promote consistency and improve accuracy	Development of data validation tools Standardised format of storage & curation of similar type data Incentivise data owners to share and collaborate	Development of improved analytical models	Provision of regular science communication training to data contributors and data consumers

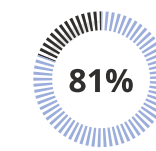
Summary of survey responses

- In a pre-workshop survey, Data Catalyst participants confirmed the data challenges that were previously identified in the TNFD Data discussion paper, as shown on the graph below:

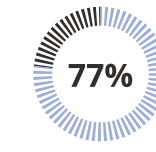


- DC participants were also asked how their tools or frameworks align with the TNFD draft disclosure recommendations. The majority of respondents stated that their tools align to the “Metrics and Targets” pillar as they provide or organize data for disclosure. This was followed by alignment to the “Strategy” pillar. No participants demonstrated alignment to the “Governance” pillar.
- Workshop participants self-declared a focus on the “Locate” and “Evaluate” phases of the LEAP framework. Within these pillars L3 and E4 were the most commonly selected sub-steps.

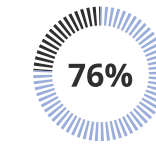
- Other insights from the survey included:



81% of DC participants are using specialized technologies



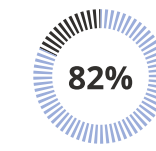
77% of DC participants indicated their data tool can assist entities in translating impact and dependency data into quantitative or qualitative risk and opportunity data



76% of DC participants indicated their data services can assist entities in choosing and measuring the performance of their responses to risks and opportunities identified, including actions taken to transition towards being nature-positive



38% of DC participants indicated that user restrictions are associated with their tools, including commercial use restrictions, re-distribution restrictions, subscriptions and licenses



82% of DC participants indicated that they engage in nature-related partnerships or collaborations, including commercial, non-commercial, academic, governmental and funding partnerships

Next steps

The initial series of Data Catalyst workshops highlighted the breadth of knowledge across the nature-related data landscape. Understandably, priorities and insights varied from participant to participant depending on their particular perspective, but feedback received suggests that the initial landscaping and group collaboration was an important first step in shaping the onward objectives of the Data Catalyst. As we set out in our Terms of Reference and have stated throughout all participant interactions, our high-level objective is to facilitate the broadest adoption of the TNFD framework in the shortest timeframe possible by addressing and if possible removing the gaps and challenges that currently exist to potential framework adopters. More specifically:

1. Convene key actors around a common objective for data on nature-related risks and opportunities, including both improving existing tools and platforms and accelerating development of new tools and platforms.
2. Establish a baseline for data relevance and credibility, while also identifying the desired future state;
3. Identify areas for innovation where required to address gaps and overcome challenges;
4. Encourage partnerships where appropriate, helping to unlock data and analytics enabling greater access;
5. Set out guidance, in the TNFD's final recommendations in September 2023, for data usage in populating metrics for the TNFD framework.

In the next phase of Data Catalyst participation, we will be focusing on facilitating broader engagement and collaboration. Additional feedback we've received from the consumers or end-users over the last two months has centred on the 'how' question of data and its usage. We therefore want to identify or create real world examples of nature-related data usage that are relevant to the TNFD framework adoption. This will help these end-users to better understand how to access and use data to address the TNFD framework, but also discover and highlight where there are existing issues/limitations/gaps etc. with data restricting assessment, disclosure or analysis.

To achieve this, there will be opportunities for the Data Catalyst members to collaborate with the **TNFD piloting program**, in which over 100 companies and financial institutions have committed to piloting the TNFD framework. As pilots report back on the issues faced, participants of the Data Catalyst will be offered the opportunity to collaborate on specific issues, leveraging their area of expertise to address these gaps and challenges. The **collaborative groups** will be sub-divided to focus on specific issues e.g. data coverage, data standardisation, transparency or accessibility. We may also look to split the workstreams up e.g. by sector using SASB's Industry Classification list as reference.

In response to numerous requests from TNFD Forum members and the wider market, we have created the first iteration of an online **Nature-related Data Tools and Analytics Catalogue** which aims to signpost for data consumers – companies and financial institutions – the tools and platforms that can be currently used in adopting the TNFD framework. The initial iteration of the catalogue should be seen as a minimum viable product building on the mapping of available tools and platforms that could currently be used in adopting the LEAP process. Please note the initial list of tools and platforms is not comprehensive but aims to provide a starting point for consumers of data, both companies and financial institutions to better understand where and how to access relevant data. The Catalogue is intended to identify the numerous entry points for adopting the TNFD framework e.g. geographically, biome specific, by sector, realm, ecosystem services or the LEAP process etc. The Catalogue will be developed iteratively based on feedback and contributions from the Data Catalyst participants and data consumer feedback and priorities. Separately, on the main TNFD website, in collaboration with the DC participants, we will **profile their organisation and services** so that their capabilities are presented.

We conducted a **Data Lab at COP15** in Montreal from December 12-15 2022, in conjunction with the near-term priority on addressing the 'how' of nature-related data usage. These were a series of eight workshops, which showcased a range of tools, platforms and approaches to data in relation to nature and nature risk, grouped by sector: financial services, agri-food, energy and utilities, forestry and construction, mining and extractive resources, marine environment, and innovation and tech.

Find out more about the Nature-related Data Catalyst, its terms of reference and its activities on the TNFD website. <https://tnfd.global/consultation-and-engagement/data-catalyst/>



www.tnfd.global