The business case for investing in biodiversity data

Fevziye Hasan[1], Jakob Nyström[2], Carina Andersson[8], André P da Silva[6], Alice Högström[9], Emma Granqvist[3], Mats Eriksson[1], Robert M Goodsell[3], Fabian Roger[10], Fredrik Ronquist[3], Tomas Roslin[7], Olga V Pettersson[5], Veronika A. Johansson[3,4], and Tobias Andermann[2,5].

[1]Museum of Evolution, Uppsala University, Norbyvägen 16, 75236, Uppsala, Sweden
[2]Department of Organismal Biology, Uppsala University, 75236, Uppsala, Sweden
[3]Department of Bioinformatics and Genetics, Swedish Museum of Natural History, Box 50007, SE-104 05, Stockholm, Sweden
[4]Global Biodiversity Information Facility (GBIF) Sweden, Stockholm, Sweden
[5]Science for Life Laboratory, Uppsala University, Uppsala, Sweden
[6]Stockholm Resilience Centre, Stockholm University, Stockholm, Sweden
[7]Swedish University of Agricultural Sciences (SLU), Department of Ecology, Ulls väg 18B, 75651, Uppsala, Sweden
[8]Knowit AB, Stockholm, Sweden
[9]Sveaskog, Wallingatan 2, SE-111 60, Stockholm, Sweden
[10]DNAir AG, Zurich, Switzerland

Fevziye Hasan ORCID: 0000-0001-9670-9775

Jakob Nyström ORCID: 0009-0008-0734-0936

Carina Andersson **-**

André P da Silva ORCID: 0000-0002-4722-8497

Alice Högström -

Emma Granqvist ORCID: 0000-0002-1513-1674

Mats Eriksson **-**

Robert M Goodsell ORCID: 0000-0002-3349-1876

Fabian Roger ORCID: 0000-0001-9865-7542

Fredrik Ronquist ORCID: 0000-0002-3929-251X

Tomas Roslin ORCID: 0000-0002-2957-4791

Olga V Pettersson ORCID: 0000-0002-5597-1870

Veronika A. Johansson ORCID: 0000-0002-3028-9947

Tobias Andermann ORCID: 0000-0002-0932-1623

Corresponding author: Fevziye Hasan; fevziye.hasan@em.uu.se

Author contributions

FH and TA conceptualised the paper and all authors participated equally in the design. FH, JN, TA developed the framework and all authors contributed critically to writing all drafts and gave final approval for publication. FH drafted the manuscript and all authors contributed to the writing.

Acknowledgements

We thank Stiftelsen Oscar och Lili Lamms Minne for funding the “Emerging Methods in Business and Biodiversity: Data is the Shared Language” workshop at Ekenäs Herrgård, Sweden, on December 5-6, 2024. The workshop was jointly organized by FH, TA, and JN from the Museum of Evolution and the Biodiversity Data Lab at the Department of Organismal Biology, Uppsala University.

Conflict of Interest

The authors have declared that there are no competing interests.

Abstract

1. The private sector is increasingly aware of its dependence on biodiversity and the financial risks and opportunities involved. This has generated a lot of demand for investing in nature-positive solutions. There is an obvious and non-negotiable basis for such initiatives: biodiversity data. Without this data and the tools built from it, no actor can assess the effects on the ecosystems they rely on. We identify two key barriers to corporate biodiversity action: (1) lack of biodiversity data and (2) challenges with biodiversity data literacy, i.e. the domain knowledge necessary to apply data products for decision making in appropriate contexts. Building on this, we present an end-to-end framework mapping biodiversity data to data products and business use cases, to establish a shared language between business and biodiversity research.

2. First, we provide examples of new technologies for generating biodiversity data at unprecedented scales, such as environmental DNA, computer vision and audio monitoring. We discuss the large amount of biodiversity data available in open databases, with a focus on the Global Biodiversity Information Facility (GBIF), including their origins, limitations, and biases. We highlight the one billion untapped primary biodiversity data points in natural history collections, and the opportunity to mobilise them into open databases using technology at relatively low cost.

3. Second, we discuss biodiversity data products, focusing on the ability to interpret, communicate, and effectively apply biodiversity models, metrics, and tools in relevant contexts. We address the challenges posed by the complexity of biodiversity, the importance of its definitions, and the use of aggregated metrics for biodiversity and ecosystem services in reporting, including the role of nature tech.

4. Third, we present the business case for investing in more and open biodiversity data, with examples of actions by companies and the finance sector. We also propose a mechanism to incentivise and reward direct investments in biodiversity data mobilisation. In conclusion, we call on businesses to prioritise financial investment in biodiversity data collection and mobilisation, to create better data products that can accelerate deployment of solutions to the biodiversity crisis.

Key words

Biodiversity data, Business and biodiversity, data mobilisation, financing biodiversity, Global Biodiversity Information Facility (GBIF), Natural History Collections, Nature Tech, Nature-based Solutions

Introduction

Biodiversity underpins essential ecosystem services that support our societies and economy (Díaz et al., 2018; Mace et al., 2012). Its rapid decline due to human activities is pushing planetary boundaries (Steffen et al., 2015), already costing the global economy over $5 trillion annually (Ranger et al., 2023). The World Economic Forum and central banks recognise the risks of biodiversity loss, but the 2024 Global Risks Report suggests these risks will only become serious in the next decade (WEF, 2024), downplaying urgency despite biodiversity loss being a current reality.

Globally, both “hard” mandatory regulations and “soft” voluntary frameworks have been introduced to address concerns about biodiversity loss. The Kunming-Montreal Global Biodiversity Framework (GBF)(COP15, 2022) has emerged as a strong influence on companies, in the same way the Paris Agreement has shaped net-zero commitments in the context of greenhouse gas-emissions (Allen et al., 2025). Target 15 of the GBF requires large companies and financial institutions to monitor, assess, and transparently disclose their biodiversity risks, dependencies, and impacts across operations, supply chains, and portfolios (COP15, 2022). Concurrently, the EU Corporate Sustainability Reporting Directive (CSRD) mandates environmental impact reporting for approximately 50,000 companies by 2025 (Faqih & Kramer, 2024). Market-led initiatives like the Taskforce on Nature-related Financial Disclosures (TNFD) are creating science-based frameworks for managing nature-related risks and opportunities, with substantial engagement from now over 500 organisations representing £17.7 trillion in assets (TNFD, 2024).The Science Based Targets Network (SBTN), originally established to guide companies in addressing the climate crisis, is now including nature and biodiversity (SBTN, 2020).

This emerging intersection of business and biodiversity has stirred both optimism (White et al., 2023) and concern (Smith et al., 2019). Among the sources of optimism is the hope that private sector interest in biodiversity could provide an opportunity to tackle the biodiversity crisis and deliver data-driven solutions to long-studied problems. However, there is also apprehension that corporate sustainability departments may adopt unscientific or oversimplified approaches, inadvertently harming the very ecosystems they aim to protect (Mair et al., 2024). For companies, integrating biodiversity into their reporting presents both challenges and opportunities. In the short term, companies must adapt quickly to comply with mandatory environmental reporting requirements (D'Amato et al., 2024). In the medium term, they face increasing pressure from investors and lenders who are focusing more and more on nature-related risks and sustainability metrics (Ingram et al., 2024). Companies that fail to meet these emerging standards may face reputational damage, higher financing costs, or reduced access to capital – whereas early adopters can gain competitive advantages (Kulionis et al., 2024).

To meet these new requirements and capitalise on opportunities, businesses are investing in Nature-based Solutions (NbS), which is an umbrella term for working with nature to benefit biodiversity and people (Seddon et al., 2020). New markets for biodiversity credits are also emerging under NbS, raising concerns due to the way carbon credits evolved, and underscoring the pressing need for scientific credibility (Aide, 2024; Swinfield et al., 2024). Simultaneously, there is rapid growth of the nature tech sector, referring to any technology that enables, accelerates, and scales businesses' nature-positive transition (Goren, 2024). Technologies that deliver "nature intelligence" analytics and tools to clients include in-situ biodiversity measurement hardware and software-as-a-service (SaaS) platforms, which ingest data from open-access biodiversity databases. The nature tech sector saw investments exceed USD 2 billion in 2022 and a compound annual growth rate of 52% since 2018 (Evison et al., 2022). Unlike the standardised CO2-equivalent metrics for carbon accounting, biodiversity's inherent complexity has resulted in widespread confusion surrounding data, metrics, reporting, and valuation approaches for businesses (Jones & Solomon, 2013).

To guide decision-making on both local and global scales, there is no substitute for reliable biodiversity data (Gerber & Iacona, 2024; Hawkins, 2024; Hobern et al., 2019; Mason Heberling et al., 2021; Musvuugwa et al., 2021). Without knowing current biodiversity trends and how specific actions impact biodiversity, we cannot make decisions that are truly "evidence-based". Despite current biases, the economic value of open biodiversity data should not be ignored, as open data on platforms like GBIF generate €3 in direct benefits for users and up to €12 in societal returns for every €1 invested, extending to business benefits (Deloitte, 2023). Despite the critical value of information and data-driven solutions, there seems to be a lack of efforts to clearly define the challenges in the biodiversity data pipeline to create suitable data-derived products for integrating business and biodiversity (Panwar et al., 2023).

The complexity of integrating biodiversity into business contexts is further heightened by its multifaceted nature, which encompasses genetic diversity, species diversity, and ecosystem diversity (CBD, 2011). While species serve as the most commonly used unit of biodiversity, impacts at one level may not reflect changes at another (Exposito-Alonso et al., 2022; Kahilainen et al., 2014). Even seemingly straightforward metrics like species richness do not always indicate better biodiversity outcomes (Hillebrand et al., 2018). Essential Biodiversity Variables (EBVs) are measurements for reporting and monitoring, developed to aid businesses in measuring biodiversity (Pereira et al., 2013), but are often based on top-down methods that rely on indirect or modeled data, which can be biased and incomplete (Granqvist et al., *In prep.*). Interpreting diversity patterns requires context, i.e. biodiversity data literacy. This complexity highlights the need for clear frameworks that can help businesses effectively integrate biodiversity considerations into their operations while maintaining scientific validity.

Accounting for biodiversity is a process that should be grounded in established scientific methods and fieldwork (Hill et al., 2005). A bottom-up approach prioritising the collection of in-situ biodiversity data at the local scale is needed and necessary to derive other EBVs (Granqvist et al., *In prep.*). Transparent and regular biodiversity data collection, including baseline measurements and ongoing monitoring, are important for scrutinising biodiversity impacts and the direction of progress in stewardship (White et al., 2023). Biodiversity data needs will vary by industry, with larger or more complex operations likely requiring a combination of data sources from field monitoring, satellite data, and open biodiversity data (see Case Study S1).

Biodiversity monitoring is valuable but insufficient without broader ecological context. Around four out of five species globally remain undiscovered, even in well-studied regions (Mora et al., 2011; Stork, 2018; Miraldo et al., 2024), highlighting critical gaps in our understanding of ecosystems. Alarmingly, despite the significant returns of investing in biodiversity data infrastructures (Deloitte, 2023), funding remains insufficient. Natural history collections and herbaria—the original biodiversity data infrastructures that house type and voucher specimens essential for confirming species identities, updating species red-list statuses, and advancing technologies such as eDNA—are closing due to short-sighted funding systems. These collections, which contain between 1.2 and 2.1 billion specimens worldwide, are inadequately represented in GBIF, with only around 200 million specimens currently included (Huybrechts et al., 2022). For plants and fungi alone, there are approximately 400 million specimens stored internationally across 3,000 herbaria, which have the potential to aid in advancing understanding of traits and predictive modeling, such as forecasting biodiversity under future climate change (Davis, 2023). The decline of these collections is a catastrophe for biodiversity, driving the silent extinction of both species and taxonomists (Löbl et al., 2023). We cannot expect to advance new technologies that collect large quantities of biodiversity data without supporting the taxonomic foundations from which they are built. Simultaneous investment towards mobilising both specimen- and observation-based primary biodiversity data is needed. Without this comprehensive, scientific approach, efforts to measure and implement biodiversity positive actions may be ineffective.

The open biodiversity data on GBIF is expansive, however, it is biased geographically and taxonomically (Troudet et al., 2017), with currently around 65% representing birds, which only make up 0.5% of all species currently known to science (GBIF, 2025). Most other groups of species remain poorly represented, leading to our knowledge of biodiversity patterns and responses being mostly restricted to a few data-rich taxonomic groups. Similar biases exist in national databases, including in best-studied regions like Finland (Roslin & Laine, 2022). New technologies offer promising solutions, such as advances in environmental DNA (eDNA) (Deiner et al., 2021), species identification using computer vision (Beery, 2023), and acoustic monitoring (Buxton et al., 2018). These innovations are revolutionising biodiversity data collection (van Klink et al., 2022; Van Klink et al., 2024), reducing the time and expertise needed for species inventories (August et al., 2015).

Using GBIF data without addressing its biases can lead to misleading conclusions (Boyd et al., 2023). This is particularly concerning for nature tech companies offering SaaS solutions that rely solely on open data for biodiversity analytics. Such practices risk leading to unsupported claims, harming biodiversity, disrupting ecosystems, compromising clients' operations, and eroding trust in biodiversity as a critical business issue. Acknowledging the local and global data biases, some initiatives are now warning companies about the “nature data gap” (Nature Tech Collective, 2024; TNFD, 2024). There is positive growing recognition that biodiversity impact reports must be validated with “ground-truthed” in-situ data collection (WWF, 2024) and enabling “nature intelligence” requires high quality biodiversity data (TNFD, 2024).

In this perspective, we turn to first principles and present an end-to-end framework (Figure 1) that charts the journey from data to business use cases. First, we explore biodiversity data, focusing on methods for collecting primary biodiversity data using new technologies from the field and natural history collections. Second, we examine biodiversity data products and the methods used to derive them, presenting a non-exhaustive set of quality markers these products should meet. Finally, we explore the business cases for biodiversity data, concentrating on impact and risk reporting, nature-based solutions, and sustainable ecosystem stewardship. We stress the need to prioritise biodiversity data collection and mobilisation, arguing that biodiversity data literacy is crucial for achieving biodiversity-positive outcomes.

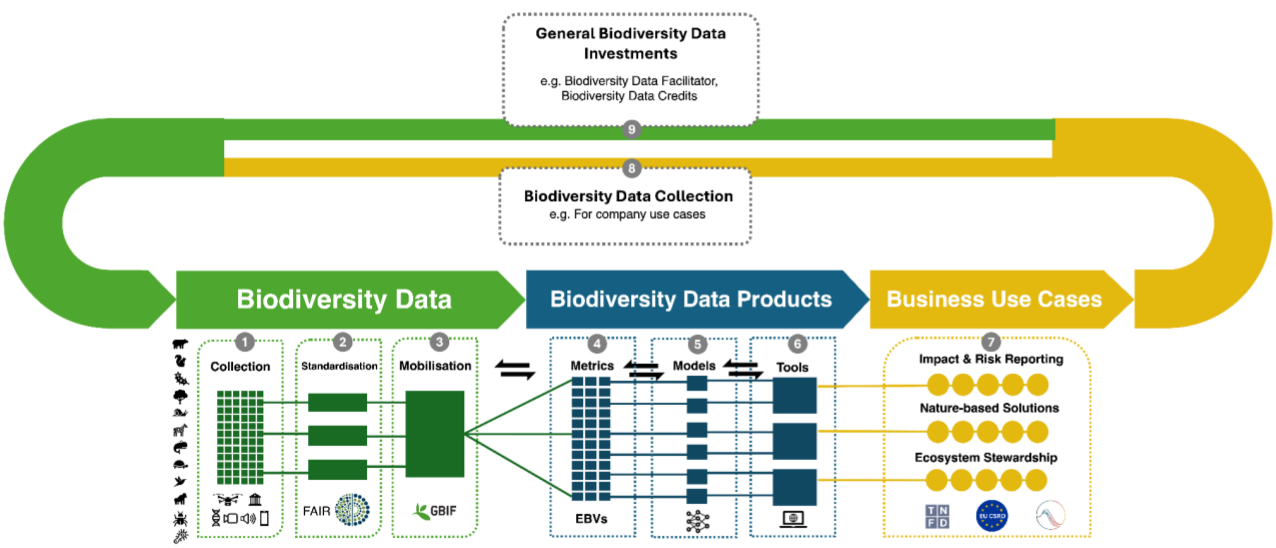

Figure 1. Framework for integrating business and biodiversity, grounded in data. This pipeline clarifies the steps to get from raw biodiversity data to business use cases. The process begins with biodiversity data (green), divided into three key components: (1) collection of primary biodiversity data from both the field and natural history collections using traditional and new technologies; (2) standardisation of this data

in alignment with FAIR data principles and the biodiversity information standards (TDWG); and (3) mobilisation of these data into the Global Biodiversity Information Facility (GBIF), the world's largest biodiversity data repository. The next step involves biodiversity data products (blue), where biodiversity experts translate raw data into (4) biodiversity metrics (e.g., EBVs); (5) data-driven and predictive models; and (6) data tools for users, which are iteratively updated as new data become available. This biodiversity expertise has traditionally been represented by academic research, consultancies, and public environmental agencies, but is increasingly adopted by the rapidly evolving nature-tech sector. Finally, business (yellow) represents the end-users of the data products (7). Important use cases for biodiversity data products include impact and risk reporting, investments in nature-based solutions (NbS) and monitoring of their outcomes, as well as better management practices through ecosystem stewardship. There are two important feedback loops in the framework. First, businesses are encouraged to invest in data collection and mobilisation of these data to the public domain, to improve their reporting and operational management (8). Second, we propose a mechanism to incentivise direct investment in biodiversity data mobilisation (9). This flow emphasises that biodiversity data is a central priority of the entire pipeline. Without high-quality biodiversity data, none of the subsequent steps are possible.

An end-to-end framework from biodiversity data to business use cases

1 Biodiversity data

The big data revolution has led to a significant increase in biodiversity data (Bayraktarov et al., 2019; Musvuugwa et al., 2021), offering opportunities to fill data gaps related to taxa and geographical distributions (Troudet et al., 2017). Primary biodiversity data, or occurrence data, constitutes the majority of data published through GBIF (GBIF, 2024) and includes three key components: taxonomic level (e.g., species, genus), location, and date (Spear et al., 2023). Collecting both observation-based and specimen-based primary biodiversity data (Figure 2) is essential to fill the biodiversity data gap, as it is foundational for quantifying abundance, understanding biodiversity patterns, mapping species distributions, assessing red list statuses, and temporal environmental change (GBIF, 2024).

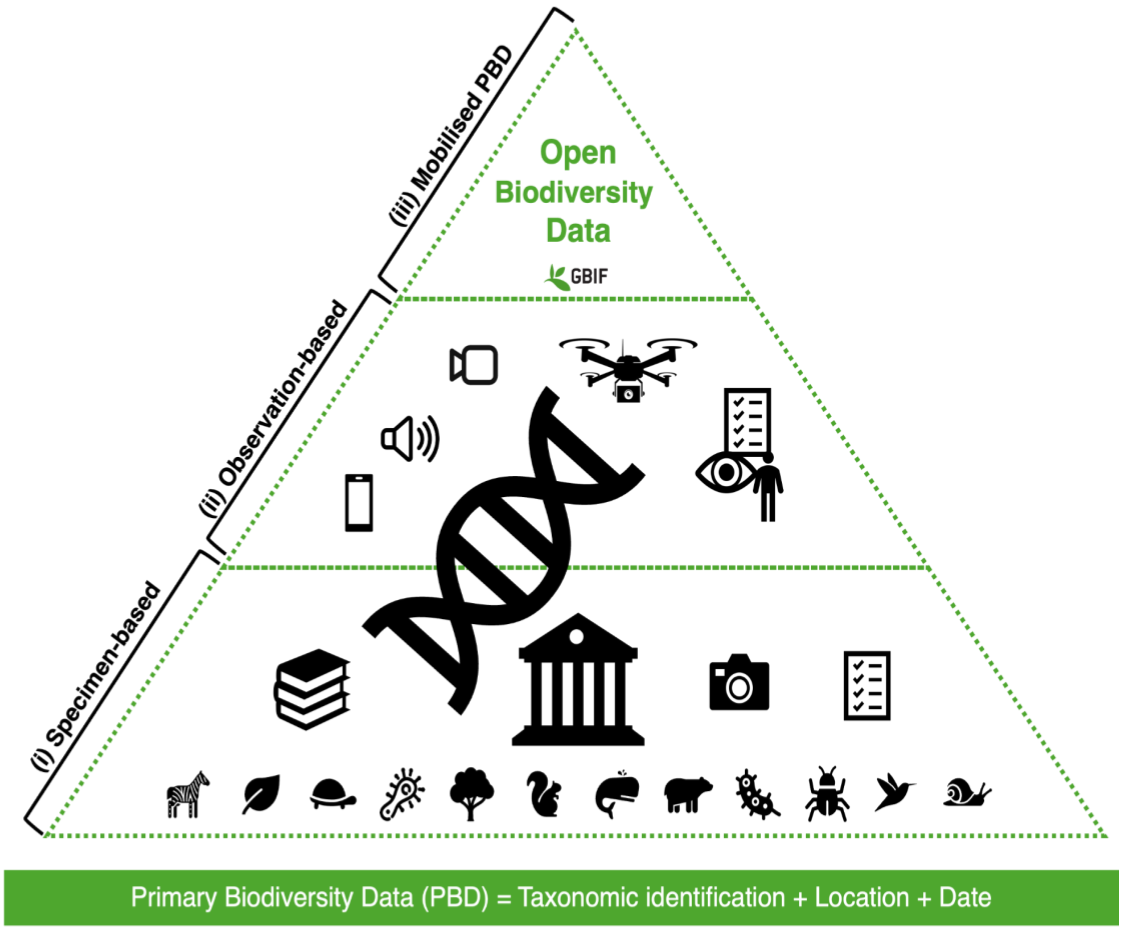

Figure 2. Pyramid diagram illustrating sources of primary biodiversity data showing that specimen-based data is foundational to observation-based data, which is foundational to integrated primary biodiversity data. (i) **Specimen-based data**: Derived from physical specimens in natural history collections, including image files, checklists, and archival materials; (ii) **Observation-based data**: Derived from traditional species inventories and technologies such as DNA methods, camera traps, audio recordings, and citizen science. DNA methods (eDNA and metabarcoding, metagenomics) overlap with both specimen and

observation data, as they require physical sample collection to generate verifiable species names; and (iii) **Mobilised primary biodiversity data**: Integrated specimen- and observation-based data on open access biodiversity datasets, such as GBIF.

New technologies offer scalable methods for collecting primary biodiversity data across ecosystems (Stephenson, 2020). For example, environmental DNA (eDNA) in combination with metabarcoding can identify species from samples like water, soil, or air (Deiner et al., 2021)(**Case Study S1**). These methods are well-suited for studying taxa with hidden diversity, such as fungi and insects, which are often challenging to access and/or difficult to identify. Using metabarcoding for biodiversity inventories has been estimated to complete the equivalent of 1,000 years of manual inventory work in just one year (Ronquist et al., 2020). Passive acoustic monitoring with autonomous sound recorders captures species sounds, monitoring birds, insects, amphibians, primates, bats, and even soil biodiversity through unique vibration patterns (Hildebrand et al., 2024; Hoefer et al., 2023; Robinson et al., 2024). AI pattern-recognition techniques analyse sounds to identify species accurately and at scale, reducing human bias and enabling simultaneous data collection across multiple sites (Buxton et al., 2018). Cameras can capture species images over time, targeting specific taxa (Bjerge et al., 2021). A global camera network, akin to meteorological weather stations, could provide real-time biodiversity data (Steenweg et al., 2017). Citizen science projects such as iNaturalist use smartphone cameras and microphones to collect biodiversity data (August et al., 2015). In addition, images and audio files collected on their platform are used as training data to advance deep learning methods, including computer vision for species identification (Beery, 2023; Høye et al., 2021). Thermal, LIDAR, hyperspectral, multispectral, and RGB sensors can be attached to drones and unmanned aerial vehicles (UAVs). Thermal drones track species by detecting heat signatures (Larsen et al., 2023). LIDAR data from stationary devices, drones, or airplanes create 3D vegetation models, while hyperspectral sensors identify tree species by analysing light reflectance patterns. RGB drones, combined with AI, can now identify plants down to species level in some cases (Mäyrä et al., 2021).

Beyond generating new data from nature, digitisation technology and AI tools are turbo-charging biodiversity data collection from natural history collections and herbaria (Nelson & Ellis, 2019). High throughput specimen digitisation, such as Angled Label Image Capture and Extraction (ALICE), uses a multi-camera setup and an associated software processing pipeline, enabling the standardisation of 800 specimens per day (Dupont & Price, 2019), and allowing data upload to GBIF for understudied and functionally important groups, such as ground beetles (Garner et al., 2024). Digitisation of 1 million herbarium specimens is estimated to take just 8 years, and hold great value for quantifying functional traits such as leaf mass per area, water-related traits, carbon fractions, and pigments, comparable to those obtained from fresh tissues (Davis, 2023). Named Entity Recognition (NER) is now used to identify and extract specimen identification, location, and date. Optical Character Recognition (OCR) is employed to convert text from images of specimen labels into machine-readable text, processing approximately 20,000 herbarium specimens at an average rate of 20 labels per hour (Takano et al., 2024). It is rare to have the opportunity to travel back in time to fill biodiversity data gaps, but we can do so by collecting data from natural history collections. These specimen-based primary biodiversity data contribute to temporal predictive modeling and link historical specimens to modern technologies, such as DNA-based species identification. Despite advances in digitisation, valuable data remains un-digitised, requiring significant mobilisation efforts (Huybrechts et al., 2022).

Consistent data standardisation is essential to ensure that biodiversity data collected from heterogeneous sources meet FAIR principles (Findable, Accessible, Interoperable, Reusable) and align with the Open Science Framework (Carroll et al., 2021; Wilkinson et al., 2016). The Biodiversity Information Standards (TDWG) develops and maintains standards for managing and sharing biodiversity data, curating and extending standards like Darwin Core (DwC). DwC ensures biodiversity data sharing by using standardised terms and vocabularies, with a namespace policy enabling universal understanding and making data machine-readable and interoperable. Extensions like the Humboldt Ecological Inventory manage ecological survey data, while the DiSSCo network standardises the aggregation and sharing of specimen-based data from natural history collections. However, big data integration faces limitations regarding metadata standards for cross-scale analysis (Maldonado et al., 2015; Wieczorek et al., 2012; Hardisty et al., 2022). Metadata standards define and manage data context, making their development vital for new technologies. For example, dnaDerivedData with MIxS (Minimum Information about any Sequence) offers guidelines on sample collection location, environmental context, DNA extraction methods, and sequencing techniques (Abarenkov et al., 2023). The Camera Trap Data Package (Camtrap DP) is a data exchange format for image data of larger animals such as mammals and birds (Bubnicki et al., 2023) currently being extended with controlled vocabularies to include broader taxonomic scope, including insects. Additionally, the Ecological Metadata Language (EML standard) provides detailed documentation of dataset characteristics for all types of biodiversity data.

Despite progress in FAIR data standardisation, challenges persist, including duplication of data points, variable quality, and interoperability issues (Pyle et al., 2021). Taxonomic changes and errors, especially in valid names and classifications, can be addressed with a mix of automation and expert curation (ChecklistBank, 2025; Whitley et al., 2024). These advancements, along with new data collection technologies, make mobilised data accessible on platforms like GBIF, supporting efforts to bridge the biodiversity data gap. Biodiversity data mobilisation involves sharing FAIR data and ensuring it is open access. Despite consensus on the value of open data in biodiversity research, motivating data collectors to digitize and share their data remains challenging. Academic incentives for data sharing include DOI citations and data paper publications, while businesses are increasingly recognising the strategic value of mobilizing biodiversity data. Businesses become data publishers by sharing their data through the GBIF Integrated Publishing Toolkit (IPT) (see Case Study S2). To date, the business sector has contributed a mere 0.3% of the published records to GBIF. Publishing through a national GBIF node's IPT is usually free of charge and enables data citation and impact tracking through assigned DOIs and UUIDs, monitored via GBIF's literature tracking system (Figure 1(8))(**Case Study S2**).

2 Biodiversity data products

There is a growing demand to transform raw biodiversity data into metrics and data products that can cater to diverse use cases and needs across different industry sectors (Burgess et al., 2024). This task requires reducing the complexity of biodiversity into manageable metrics, which arguably is an exercise of great oversimplification, yet a necessary one. With this inherent constraint in mind, we reflect on several issues in the current state-of-the-art of biodiversity reporting and the underlying data-products.

In the context of biodiversity impact reporting, data products that provide regional or global heatmaps of biodiversity metrics are in high demand, as they allow easy area-based calculation of biodiversity value and impact. One example of such a data product is the Biodiversity Intactness Index (BII, Newbold et al., 2015; Phillips et al., 2021), which is proposed as a component indicator in the COP 16 draft of the GBF monitoring framework (CBD, 2024). Another example of a biodiversity model used in business context is GLOBIO (Schipper et al, 2020), also proposed as a GBF indicator (CBD, 2024). However, many of such global heat maps generated (Myers et al., 2000) are only weakly linked to the evaluation of the biodiversity impact of *specific decisions and actions*. For biodiversity data products to be actionable in a corporate setting, they need to relate biodiversity impacts and risks to operational and financial decisions taken by companies, so that impact tradeoff analysis can be performed. Examples include spatial planning for forest and agricultural land management, deciding from which countries and regions to source materials and products, and investments into new factories and logistics facilities. A common denominator for many use cases is the urgent need for regional and local data and models (as opposed to global) to ensure high-quality analysis and drawing the right conclusions.

While the BII and other similar data-products are being used for company impact assessment and reporting, a concern raised is that the underlying models are largely untested for their predictive performance and their agreement with other indicators of biodiversity impact (Martin et al., 2019, Nyström, 2024). We see a big risk that insufficiently tested data products provide the foundations for company impact reporting and nature investments, with potentially negative consequences. This problem is further exacerbated by the quickly developing nature tech market, driven by the demand for attractive, ready-to-use biodiversity solutions and data. The absence of a thorough quality-checking and peer-review process in this context lends reason for concern and makes it difficult for customers to distinguish between "snake-oil salesmen" with questionable data products and those built on solid foundations. However, as outlined above, even models and data products that have been reviewed by the academic peer-review process, risk being mis-applied for purposes they were not designed for. Part of the reason for this misapplication is a lack of guidance on the use of existing and emerging biodiversity data products and metrics, which we identify in this article as the challenge of **biodiversity data literacy**.

To provide guidance to help businesses and other stakeholders to identify high-quality data products, which are often based on statistical/machine learning approaches, we have compiled a non-exhaustive set of quality markers that such products should exhibit. These recommendations broadly apply to both data products provided by academic groups and non-profit organisations, as well as nature tech solutions.

- **Out of sample testing and ground truthing:** Predictive models should always be tested on data not used for model training, to approximate its performance when used in real-world applications. This tests the model's ability to generalise patterns beyond its training data. Ideally, when making predictions on smaller scales, new measurement data can be collected and used for testing (ground-truthing). In cases where new measurements are not possible, e.g. when working on very large scales, cross-validation can be used to simulate the application of the model on new data. This involves repeatedly splitting the data into training and test sets (cross-validation folds), and evaluating the performance on each test fold. Importantly, the evaluation should be based on the relevant prediction task to be tackled rather than generic considerations (Abrego & Ovaskainen, 2023). It should also account for spatial and environmental dependencies in the data (Roberts et al, 2017). Such a cross-validation approach, if done correctly, will in principle provide similar indications of performance as “real” ground-truthing.

- **Local inference**: The location of the training data should be disclosed. This is important as inferences should typically only be made within the given region or set of regions where the model was trained. Inferences outside of the training area can be highly problematic as they might miss different parts of the environmental, anthropogenic, and geographic variable space potentially leading to erroneous predictions.

- **Uncertainty quantification:** Any data product based on model predictions should address the issue of uncertainty. Each predicted value needs to have an associated uncertainty measure, expressing how confident the model is in the prediction. This allows users to filter output data based on a required confidence threshold and also addresses the previously mentioned issue with spatial biases in the training data. For instance, regions that are poorly represented in the training data tend to be associated with higher uncertainty. Data products without accessible and transparent uncertainty estimation give a false sense of precision that is detrimental, and sometimes dangerous, for decision-making.

- **Transparency and limitations:** All biodiversity data products should have a clear list of limitations to inform users about intended purpose, appropriate and non-recommended use cases, limitations of the underlying data in terms of taxonomic and spatial biases, as well as known cases or areas of poor performance and high uncertainty. The underlying training data should be accessible with an open license. While we acknowledge that nature tech companies need to safeguard code and certain data for competitive reasons, we strongly argue that any data products of academic or non-profit origin should have their data and code repositories publicly available. Without such access, there can be no real peer-review process. Nature tech solutions that rely on underlying scientific models should clearly disclose the sources (academic papers and code) and should ideally include a high-level documentation of the overall quantitative approach used.

Despite best efforts regarding methodological considerations and quality assessments, data products are only as reliable as the data they are derived from. At present, the biggest bottleneck to better biodiversity models is arguably the lack of contextualised data, particularly in view of the vast taxonomic and spatial biases that exist. Closing the biodiversity data gaps is critical for enhancing the accuracy and reliability of biodiversity metrics. We argue that this can be achieved through significant financial investment in data collection and mobilisation. In the next section, we outline ways to incentivise companies to invest in large-scale biodiversity data generation and mobilisation, to support high-impact use cases.

3 Business use cases and incentives for investing in biodiversity data

Improved impact and risk reporting

Businesses are facing mounting pressure to assess and disclose their biodiversity impacts and risks, from mandatory compliance under frameworks like the CSRD to voluntary reporting initiatives such as TNFD, which are becoming standard expectations for investors and lenders. With regulations like these, businesses must prepare to assess their biodiversity impacts and risks, requiring long-term investment towards biodiversity data. To meet compliance requirements for CSRD, companies set measurable, science-based biodiversity targets to track and improve their impact (ESRS E4, 2023). Some assessments will be needed at different scales, such as biodiversity impact assessments at the product level (life cycle analysis), requiring meaningful, decision-useful data that can withstand scrutiny from regulators, investors, and stakeholders. As we have outlined here, the biodiversity data gap and biases in GBIF reduce the ability for meaningful inference and puts operations and compliance at risk. This was further highlighted in TNFD's 2023 scoping study, highlighting that the available "nature data" is outdated, inconsistent, and lacks the resolution required to inform confident decision-making (TNFD, 2023). With biodiversity data accessibility, quality, comparability, verifiability and assurability being key concerns of market participants, the TNFD proposes testing the efficacy of a "Nature Data Public Facility" to provide accessible, decision-useful nature data for corporate decision-making (TNFD, 2024), which, if approved, will continue to ingest biodiversity data from GBIF.

Companies that proactively collect ground-truthed biodiversity data and publish it through GBIF can enhance their transparency and gain a competitive edge, as it serves as a compliance indicator for GBF Target 19 under the CBD (Figueira et al., 2023). Using GBIF metrics as KPIs (key performance indicators) in sustainability reports demonstrates a clear commitment to biodiversity disclosure, signaling accountability to investors and other stakeholders. For instance, TotalEnergies became a GBIF data publisher in 2019, sharing over 51,000 biodiversity records across three continents (see Case Study S2 – Figure S2). Publishing through GBIF increases reliability, corporate visibility and compliance reporting, as companies can use their published data to track progress against biodiversity targets and meet regulatory expectations.

Both the CSRD and TNFD frameworks highlight the need for accessible, high-quality biodiversity data. Since TNFD's proposed infrastructure will likely rely on platforms like GBIF, businesses and other investors publishing data on GBIF are not only meeting immediate compliance requirements but also build a temporal record of biodiversity action, preparing for future voluntary reporting needs. In the framework presented here (Figure 1), we outline the steps that can be taken to improve biodiversity data coverage on GBIF, highlighting data collection and mobilisation as priority areas where businesses can take impactful action.

Nature-based solutions and sustainable ecosystem stewardship

NbS includes a broad suite of interventions with nature-positive outcomes for businesses to demonstrate sustainable ecosystem stewardship both within company supply chains and through making biodiversity positive investments towards projects. The NbS interventions for sectors that directly depend on ecosystems within their supply chains increasingly use new technologies for continuous data collection and real-time analysis. This monitors biodiversity impacts and dependencies within the supply chain to aid in informed decision-making for sustainable practices. For example, regenerative farming practices must measure their biodiversity over time to claim biodiversity positive impacts on soil biodiversity, nutrient cycling, carbon sequestration, and water retention through practices like crop rotation and no-till farming (Case Study 1). For companies extracting natural resources, baseline biodiversity surveys prior to business operations are needed to quantify the change in biodiversity and assess the true environmental impact (Case study 2). More examples include ensuring marine biodiversity data are collected to identify needs for selective fishing; agroforestry to improve habitat heterogeneity in agricultural land to increase biodiversity and monitoring functionally important biodiversity, such as pollinators, introducing interventions to enhance ecosystem services through planting native flowers, reducing pesticide use to boost pollination and crop yields. Baseline biodiversity data and continuous monitoring within the supply chain enable long-term stewardship of natural capital by assessing the impacts of various management practices and forecasting possible risks associated with biodiversity loss, such as reduced carbon sequestration (Case Study 1).

For NbS investments in biodiversity-positive projects beyond business supply chains, including market-based instruments like biodiversity credits, offsets, subsidies, tradable permits, and payments for ecosystem services, which are continually evolving, biodiversity data is essential for monitoring, reporting, and verifying project impacts and contributions to natural capital. A data-driven "Internet of Things" for ecosystems could enable comprehensive monitoring and verification across taxa and environmental variables, ensuring the credibility and effectiveness of biodiversity-positive projects. Initiatives like mangrove restoration, biodiversity-friendly infrastructure (e.g., wildlife corridors), rewilding through native species reintroduction, and seagrass meadow restoration must prioritise monitoring biodiversity and ecosystem variables to ensure logical conclusions of its value. Requiring these projects to mobilise their biodiversity monitoring data to GBIF ensures accountability and reliability. For companies already engaged in NbS projects, including biodiversity credits, mobilising data to GBIF provides an opportunity to create publicly available biodiversity datasets, enhance transparency in their environmental efforts, and contribute to global biodiversity knowledge.

With incentives for investing in NbS through supply chains and credit systems, and the recognition of the need to address the "nature data gap", businesses and investors have a significant opportunity to require evidence of biodiversity data collection and mobilisation to open platforms like GBIF, showcasing responsibility in their portfolios while advancing meaningful ecosystem stewardship.

Biodiversity data certification

Building on the business use cases and incentives for investing in biodiversity data, we propose biodiversity data certification as a structured approach to quantify, validate, and showcase contributions to biodiversity data collection and mobilisation for quantifiable long-term impact (Figure 3). Many action-oriented initiatives, such as Business For Nature and Finance for Biodiversity, have emerged to raise awareness and secure private sector commitments towards financing biodiversity. In the end-to-end framework presented here (Figure 1), we propose two key feedback loops to support sustainable financing for primary biodiversity data collection and mobilisation from data providers. The first involves businesses collecting and mobilising their own data (Figure 1(8)); the second involves a biodiversity data facilitator, such as a not-for-profit organisation, helping to channel investment from businesses and other stakeholders to address data gaps and biases in global open biodiversity databases (Figure 1(9)).

This proposed approach of an independent biodiversity data facilitator aligns financial investment with the goals of the GBF, offering a novel way to fund and manage biodiversity data for long-term impact. Businesses and other investors receive biodiversity data certifications, while data providers receive financing for contributing data to the public domain. The economic, societal, and business returns on investment in biodiversity data are manifold. However, there are currently no innovative financial incentives to collect and mobilise primary biodiversity data at scale. Sustainable financing is needed to support the collection, standardisation, and mobilisation of biodiversity data from both specimen-based and observation-based sources (Figure 2). Economic investment and return on investment can be tracked by the number of biodiversity data points shared with platforms like GBIF. The facilitator could apply domain knowledge to enable to target funding for data collection to fill the most important gaps in global biodiversity databases, for example by focusing on understudied taxa, such as fungi or insects, which could have a significant impact on addressing the global biodiversity data gap by creating sustainable financial flows for data providers. Mobilising data onto GBIF ensures credibility, and that biodiversity information is available as a public good—accessible and regularly updated—serving as a valuable resource for businesses, policymakers, and conservation efforts (Figure 3).

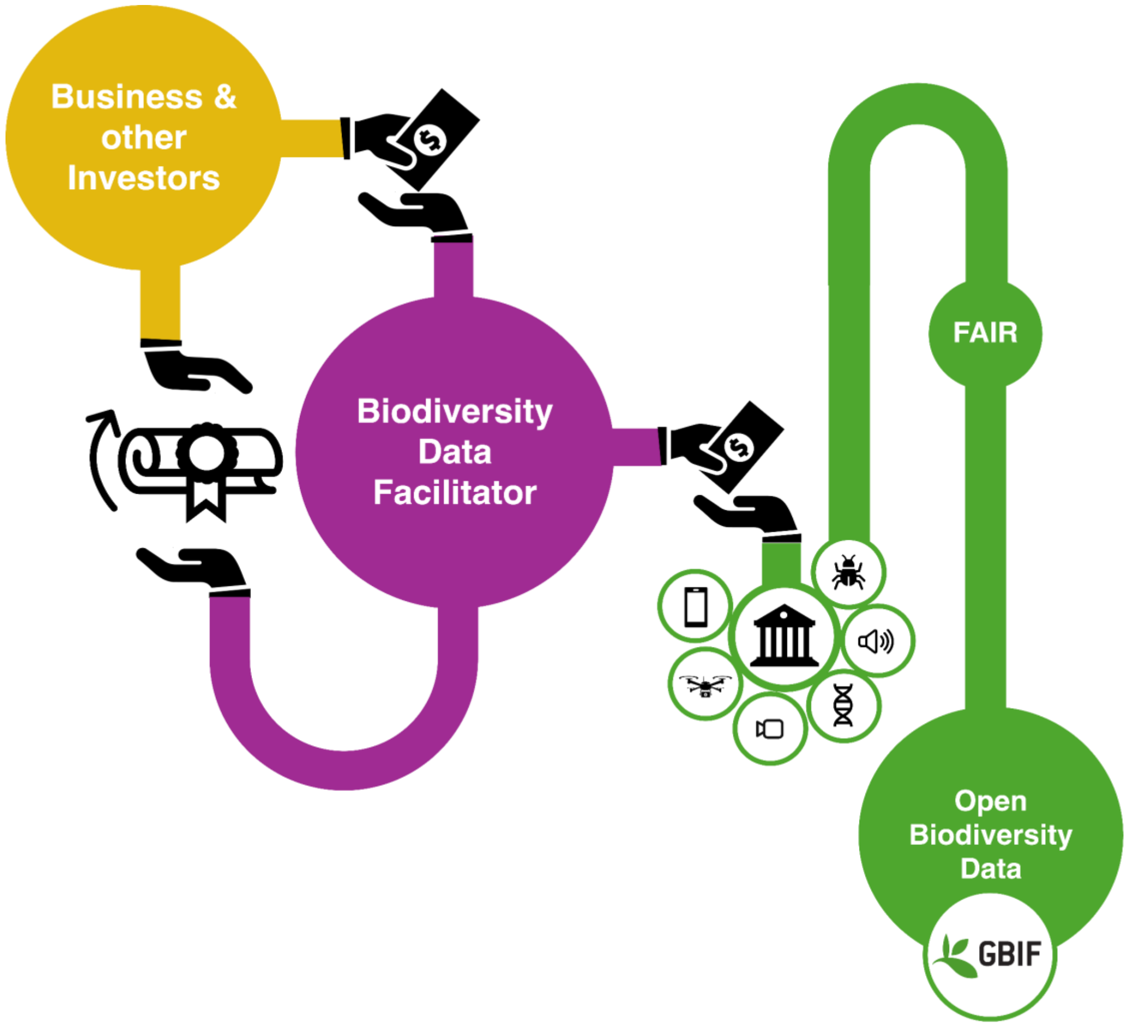

Figure 3. Innovative financing model for biodiversity data via a not-for-profit biodiversity data facilitator. Businesses and other investors (yellow) make financial investments towards data mobilisation through the biodiversity data facilitator (purple), which channels this investment into funds for partners that collect biodiversity data, advance data standards and mobilise biodiversity data onto GBIF. In return, businesses receive certification recognising their positive biodiversity actions.

Summary and call-to-action

The private sector faces an urgent challenge: understanding and managing biodiversity risks that directly impact business operations. While companies increasingly recognize their dependence on nature, two critical barriers persist - insufficient biodiversity data and limited data literacy. Recognising that positive biodiversity actions take time to show effects, biodiversity data collection and mobilisation should be prioritised. New technologies mean that we have the opportunity to prioritise the mobilisation of both specimen- and observation-based primary biodiversity data. New technologies like environmental DNA and computer vision are providing unprecedented opportunities for data collection at scale, while vast untapped resources exist in natural history collections that have great potential in advancing predictive models. GBIF provides extensive open data, but businesses need better tools to interpret and apply this information effectively. Creating meaningful biodiversity data products requires bridging the gap between scientific expertise and corporate users through reliable metrics and models grounded in biodiversity data. We call on businesses and stakeholders to urgently invest in biodiversity data collection and mobilisation towards open data infrastructures like GBIF. Such investment is crucial for developing effective biodiversity models, metrics, and data products that enable informed decision-making and support biodiversity conservation efforts. With proper investment in data collection and analysis, companies can better assess their environmental impact, manage risks, and have a positive impact on biodiversity.

References

Abarenkov, K., Andersson, A. F., Bissett, A., Finstad, A. G., Fossøy, F., Grosjean, M., Hope, M., Jeppesen, T. S., Kõljalg, U., Lundin, D., Nilsson, R. H., Prager, M., Provoost, P., Dmitry, Saara, S., Cecilie, S., Tobias, S., & Frøslev, G. (2023). *Publishing DNA-derived data through biodiversity data platforms*. https://doi.org/10.35035/doc-vf1a-nr22

Aide, T. M. (2024). The Biodiversity Credit Market needs rigorous baseline, monitoring, and validation practices. *Npj Biodiversity 2024 3:1, 3*(1), 1–4. https://doi.org/10.1038/s44185-024-00062-6

Allen, M. R., Friedlingstein, P., Girardin, C. A. J., Jenkins, S., Malhi, Y., Mitchell-Larson, E., Peters, G. P., & Rajamani, L. (2025). *Annual Review of Environment and Resources Net Zero: Science, Origins, and Implications*. 58. https://doi.org/10.1146/annurev-environ-112320

August, T., Harvey, M., Lightfoot, P., Kilbey, D., Papadopoulos, T., & Jepson, P. (2015). Emerging technologies for biological recording. *Biological Journal of the Linnean Society*, *115*(3), 731–749. https://doi.org/10.1111/BIJ.12534

Bayraktarov, E., Ehmke, G., O'Connor, J., Burns, E. L., Nguyen, H. A., McRae, L., Possingham, H. P., & Lindenmayer, D. B. (2019). Do big unstructured biodiversity data mean more knowledge? *Frontiers in Ecology and Evolution*, *7*(JAN), 426167. https://doi.org/10.3389/FEVO.2018.00239/BIBTEX

Beery, S. M. (2023). *Where the Wild Things Are: Computer Vision for Global-Scale Biodiversity Monitoring*.

Burgess, N. D., Ali, N., Bedford, J., Bhola, N., Brooks, S., Cierna, A., Correa, R., Harris, M., Hargey, A., Hughes, J., McDermott-Long, O., Miles, L., Ravilious, C., Rodrigues, A. R., Soesbergen, A. van, Sihvonen, H., Seager, A., Swindell, L., Vukelic, M., … Butchart, S. H. M. (2024). Global Metrics for Terrestrial Biodiversity. *Annual Review of Environment and Resources*, *49*(Volume 49, 2024), 673–709. https://doi.org/10.1146/ANNUREV-ENVIRON-121522-045106

Bjerge, K., Nielsen, J. B., Sepstrup, M. V., Helsing-Nielsen, F., & Høye, T. T. (2021). *An Automated Light Trap to Monitor Moths (Lepidoptera) Using Computer Vision-Based Tracking and Deep Learning*. https://doi.org/10.3390/s21020343

Boyd, R. J., Powney, G. D., & Pescott, O. L. (2023). We need to talk about nonprobability samples. *Trends in Ecology & Evolution*, *38*(6), 521–531. https://doi.org/10.1016/J.TREE.2023.01.001

Buxton, R. T., McKenna, M. F., Clapp, M., Meyer, E., Stabenau, E., Angeloni, L. M., Crooks, K., & Wittemyer, G. (2018). Efficacy of extracting indices from large-scale acoustic recordings to monitor biodiversity. *Conservation Biology*, *32*(5), 1174–1184. https://doi.org/10.1111/COBI.13119

Bubnicki, J. W., Norton, B., Baskauf, S. J., Bruce, T., Cagnacci, F., Casaer, J., Churski, M., Cromsigt, J. P. G. M., Farra, S. D., Fiderer, C., Forrester, T. D., Hendry, H., Heurich, M., Hofmeester, T. R., Jansen, P. A., Kays, R., Kuijper, D. P. J., Liefting, Y., Linnell, J. D. C., … Desmet, P. (2024). Camtrap DP: an open standard for the FAIR exchange and archiving of camera trap data. *Remote Sensing in Ecology and Conservation*, *10*(3), 283–295. https://doi.org/10.1002/RSE2.374

Carroll, S. R., Herczog, E., Hudson, M., Russell, K., & Stall, S. (2021). Operationalizing the CARE and FAIR Principles for Indigenous data futures. In *Scientific Data* (Vol. 8, Issue 1). Nature Research. https://doi.org/10.1038/s41597-021-00892-0

CBD. (2011). *Convention on Biological Diversity : text and annexes / Secretariat of the Convention on Biological Diversity*.

CBD. (2024). *Monitoring framework for the Kunming-Montreal Global Biodiversity Framework: Sixteenth meeting Cali, Colombia, 21 October–1 November 2024 Agenda item 10 Mechanisms for planning, monitoring, reporting and review*. https://www.cbd.int/doc/c/5044/ea79/105d29801a3efae8df742c93/cop-16-l-26-en.pdf

ChecklistBank. (2025). https://www.checklistbank.org/

COP15. (2022). *The Convention on Biological Diversity Kunming-Montreal Global Biodiversity Framework 15/4*. https://www.cbd.int/conferences/post20202CBD/WG8J/11/7,CBD/SBSTTA/23/9,CBD/SBSTTA/24/12andCBD/SBI/3/21,respectively.

Davis, C. C. (2023). The herbarium of the future. *Trends in Ecology & Evolution*, *38*(5), 412–423. https://doi.org/10.1016/J.TREE.2022.11.015

D'Amato, D., La Notte, A., Damiani, M., & Sala, S. (2024). Biodiversity and ecosystem services in business sustainability: Toward systematic, value chain-wide monitoring that aligns with public accounting. *Journal of Industrial Ecology*, *28*(5), 1030–1044. https://doi.org/10.1111/JIEC.13521

Dasgupta, P. (2021). *The Economics of Biodiversity: The Dasgupta Review*.

Deloitte. (2023). *Economic valuation and assessment of the impact of the GBIF network*. https://www.deloitte.com/au/en/services/economics/perspectives/total-economic-value-open-access-database-living-world.html

Deiner, K., Yamanaka, H., & Bernatchez, L. (2021). The future of biodiversity monitoring and conservation utilizing environmental DNA. *Environmental DNA*, *3*(1), 3–7. https://doi.org/10.1002/EDN3.178

Díaz, S., Pascual, U., Stenseke, M., Martín-López, B., Watson, R. T., Molnár, Z., Hill, R., Chan, K. M. A., Baste, I. A., Brauman, K. A., Polasky, S., Church, A., Lonsdale, M., Larigauderie, A., Leadley, P. W., Van Oudenhoven, A. P. E., Van Der Plaat, F., Schröter, M., Lavorel, S., … Shirayama, Y. (2018). Assessing nature's contributions to people: Recognizing culture, and diverse sources of knowledge, can improve assessments. In *Science* (Vol. 359, Issue 6373, pp. 270–272). American Association for the Advancement of Science. https://doi.org/10.1126/science.aap8826

Engel, M. S., Ceríaco, L. M. P., Daniel, G. M., Dellapé, P. M., Löbl, I., Marinov, M., Reis, R. E., Young, M. T., Dubois, A., Agarwal, I., Lehmann, P. A., Alvarado, M., Alvarez, N., Andreone, F., Araujo-Vieira, K., Ascher, J. S., Baêta, D., Baldo, D., Bandeira, S. A., … Zacharie, C. K. (2021). The taxonomic impediment: a shortage of taxonomists, not the lack of technical approaches. *Zoological Journal of the Linnean Society*, *193*(2), 381–387. https://doi.org/10.1093/ZOOLINNEAN/ZLAB072

ESRS E4. (2023). *Annex 1 - ESRS-E4: European Sustainability Reporting Standards Environmental standard No.4 - Biodiversity and Ecosystems, Delegated Act.* https://eur-lex.europa.eu/legal-content/en/ALL/?uri=CELEX:32023R2772#anx_I

Evison, W., Gill, E., Moussa, T., & O'Brien, D. (2022). *A surge in nature tech investing | PwC*. https://www.pwc.com/gx/en/services/sustainability/publications/surge-in-nature-tech-investing.html

Exposito-Alonso, M., Booker, T. R., Czech, L., Gillespie, L., Hateley, S., Kyriazis, C. C., Lang, P. L. M., Leventhal, L., Nogues-Bravo, D., Pagowski, V., Ruffley, M., Spence, J. P., Toro Arana, S. E., WeiÃŸ, C. L., & Zess, E. (2022). Genetic diversity loss in the Anthropocene. *Science (New York, N.Y.)*, *377*(6613), 1431–1435. https://doi.org/10.1126/SCIENCE.ABN5642

Faqih, A., & Kramer, R. (2024). The Corporate Sustainability Reporting Directive (CSRD): An Effective Approach to Implementation. *DIPONEGORO JOURNAL OF ACCOUNTING*, *13*(4), 1–9. http://ejournal-s1.undip.ac.id/index.php/accounting

Figueira, R., Beja, P., Villaverde, C., Vega, M., Cezón, K., Messina, T., Archambeau, A.-S., Rukaya, Dag, J., & Dairo Escobar, E., (2023). *Guidance for private companies to become data publishers through GBIF*. https://doi.org/10.35035/doc-b8hq-me03

FinBio. (2023). *FinBio annual report 2023*. https://finbio.org/wp-content/uploads/2024/04/FinBio-annual-report-2023.pdf

Garner, B. H., Creedy, T. J., Allan, E. L., Crowther, R., Devenish, E., Kokkini, P., Livermore, L., Lohonya, K., Lowndes, N., Wing, P., & Vogler, A. P. (2024). The taxonomic composition and chronology of a museum collection of Coleoptera revealed through large-scale digitisation. Frontiers in Ecology and Evolution, 12. https://doi.org/10.3389/fevo.2024.1305931

GBIF. (2024). https://www.gbif.org/

GBIF. (2024). *GBIF Science Review No.11*. https://www.gbif.org/document/5N9YVBkTP3y7kqhFQviowM/gbif-science-review-no-11

Gerber, L. R., & Iacona, G. D. (2024). Aligning data with decisions to address the biodiversity crisis. *PLOS Biology*, *22*(6), e3002683. https://doi.org/10.1371/JOURNAL.PBIO.3002683

Goren, G. (2024). *Developing the Nature Tech Taxonomy Framework*. https://www.naturetechcollective.org/stories/nature-tech-taxonomy-framework#Download%20Whitepaper

Gonzalez, A., Chase, J. M., & O'Connor, M. I. (2023). A framework for the detection and attribution of biodiversity change. Philosophical Transactions of the Royal Society B, 378(1881). https://doi.org/10.1098/RSTB.2022.0182

Granqvist, E., Goodsell, R. M., & Ronquist, F. (*In prep*). The transformative potential of eDNA-based biodiversity. *Current Opinion in Environmental Sustainability*.

Hawkins, F. (2024). How will better data (and better use of data) enable us to save the planet? *PLOS Biology*, *22*(6), e3002689. https://doi.org/10.1371/JOURNAL.PBIO.3002689

Hildebrand, J., Wiggins, S., Baumann-Pickering, S., Frasier, K., & Roch, M. A. (2024). The past, present and future of underwater passive acoustic monitoring. *The Journal of the Acoustical Society of America*, *155*(3_Supplement), A96–A96. https://doi.org/10.1121/10.0026934

Hillebrand, H., Blasius, B., Borer, E. T., Chase, J. M., Downing, J. A., Eriksson, B. K., Filstrup, C. T., Harpole, W. S., Hodapp, D., Larsen, S., Lewandowska, A. M., Seabloom, E. W., Van de Waal, D. B., & Ryabov,

A. B. (2018). Biodiversity change is uncoupled from species richness trends: Consequences for conservation and monitoring. *Journal of Applied Ecology*, *55*(1), 169–184. https://doi.org/10.1111/1365-2664.12959

Hill D, Fasham M, Tucker G, Shewry M & Shaw P. (2005). Handbook of Biodiversity Methods: Survey, Evaluation and Monitoring. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511542084

Hobern, D., Baptiste, B., Copas, K., Guralnick, R., Hahn, A., van Huis, E., Kim, E. S., McGeoch, M., Naicker, I., Navarro, L., Noesgaard, D., Price, M., Rodrigues, A., Schigel, D., Sheffield, C. A., & Wieczorek, J. (2019). Connecting data and expertise: A new alliance for biodiversity knowledge. *Biodiversity Data Journal*, *7*. https://doi.org/10.3897/BDJ.7.e33679

Hoefer, S., McKnight, D. T., Allen-Ankins, S., Nordberg, E. J., & Schwarzkopf, L. (2023). Passive acoustic monitoring in terrestrial vertebrates: a review. In *Bioacoustics* (Vol. 32, Issue 5, pp. 506–531). Taylor and Francis Ltd. https://doi.org/10.1080/09524622.2023.2209052

Høye, T. T., Ärje, J., Bjerge, K., Hansen, O. L. P., Iosifidis, A., Leese, F., Mann, H. M. R., Meissner, K., Melvad, C., & Raitoharju, J. (2021). Deep learning and computer vision will transform entomology. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(2), e2002545117. https://doi.org/10.1073/PNAS.2002545117/-/DCSUPPLEMENTAL

Hughes, A. C., Orr, M. C., Ma, K., Costello, M. J., Waller, J., Provoost, P., Yang, Q., Zhu, C., & Qiao, H. (2021). Sampling biases shape our view of the natural world. *Ecography*, *44*(9), 1259–1269. https://doi.org/10.1111/ECOG.05926

Huybrechts, P., Trekels, M., & Groom, Q. (2022). How Much of Biodiversity is Represented in Collections: A big data workflow of aggregated occurrence data. *Biodiversity Information Science and Standards 6: E94279*, *6*, e94279-. https://doi.org/10.3897/BISS.6.94279

Ingram, J. C., McKenzie, E. J., Bagstad, K. J., Finisdore, J., van den Berg, R., Fenichel, E., Vardon, M., Posner, S., Santamaria, M., Mandle, L., Barker, R., & Spurgeon, J. (2024). Leveraging natural capital accounting to support businesses with nature-related risk assessments and disclosures. *Philosophical Transactions of the Royal Society B*, *379*(1903). https://doi.org/10.1098/RSTB.2022.0328

Jones, M. J., & Solomon, J. F. (2013). Problematising accounting for biodiversity. *Auditing & Accountability Journal*, *26*(5), 668–687. https://doi.org/10.1108/AAAJ-03-2013-1255

Kahilainen, A., Puurtinen, M., & Kotiaho, J. S. (2014). Conservation implications of species–genetic diversity correlations. *Global Ecology and Conservation*, *2*, 315–323. https://doi.org/10.1016/J.GECCO.2014.10.013

Kulionis, V., Pfister, S., & Fernandez, J. (2024). Biodiversity impact assessment for finance. *Journal of Industrial Ecology*, *28*(5), 1321–1335. https://doi.org/10.1111/JIEC.13515

Larsen, H. L., Møller-Lassesen, K., Enevoldsen, E. M. E., Madsen, S. B., Obsen, M. T., Povlsen, P., Bruhn, D., Pertoldi, C., & Pagh, S. (2023). Drone with Mounted Thermal Infrared Cameras for Monitoring Terrestrial Mammals. *Drones*, *7*(11). https://doi.org/10.3390/drones7110680

Löbl, I., Klausnitzer, B., Hartmann, M., & Krell, F.-T. (2023). *The Silent Extinction of Species and Taxonomists-An Appeal to Science Policymakers and Legislators*. https://doi.org/10.3390/d15101053

Liu, D. (2024). We must train specialists in botany and zoology — or risk more devastating extinctions. *Nature*, *633*(8031), 741. https://doi.org/10.1038/D41586-024-03072-3

Mace, G. M., Norris, K., & Fitter, A. H. (2012). Biodiversity and ecosystem services: A multilayered relationship. *Trends in Ecology and Evolution*, *27*(1), 19–26. https://doi.org/10.1016/J.TREE.2011.08.006/ASSET/0F6FA3A0-5727-4F75-8747-D6780EB3193C/MAIN.ASSETS/FX2.GIF

Mair, L., Elnahass, M., Xiang, E., Hawkins, F., Siikamaki, J., Hillis, L., Barrie, S., & McGowan, P. J. K. (2024). Corporate disclosures need a biodiversity outcome focus and regulatory backing to deliver global conservation goals. *Conservation Letters*, *17*(4), e13024. https://doi.org/10.1111/CONL.13024

Maldonado, C., Molina, C. I., Zizka, A., Persson, C., Taylor, C. M., Albán, J., Chilquillo, E., Rønsted, N., & Antonelli, A. (2015). Estimating species diversity and distribution in the era of Big Data: To what extent can we trust public databases? *Global Ecology and Biogeography*, *24*(8), 973–984. https://doi.org/10.1111/GEB.12326/SUPPINFO

Mason Heberling, J., Miller, J. T., Noesgaard, D., Weingart C , S. B., Schigel, D., & Soltis, D. E. (2021). *Data integration enables global biodiversity synthesis*. https://doi.org/10.1073/pnas.2018093118/-/DCSupplemental

Musvuugwa, T., Gladmond Dlomu, M., Adebowale, A., Mahmoud, M., & Antonio Gutiérrez, P. (2021). Big Data in Biodiversity Science: A Framework for Engagement. *Technologies 2021, Vol. 9, Page 60*, *9*(3), 60. https://doi.org/10.3390/TECHNOLOGIES9030060

Mäyrä, J., Keski-Saari, S., Kivinen, S., Tanhuanpää, T., Hurskainen, P., Kullberg, P., Poikolainen, L., Viinikka, A., Tuominen, S., Kumpula, T., & Vihervaara, P. (2021). Tree species classification from airborne hyperspectral and LiDAR data using 3D convolutional neural networks. *Remote Sensing of Environment*, *256*. https://doi.org/10.1016/j.rse.2021.112322

Martin, P. A., Green, R. E., & Balmford, A. (2019). The biodiversity intactness index may underestimate losses. *Nature Ecology & Evolution 2019 3:6*, *3*(6), 862–863. https://doi.org/10.1038/s41559-019-0895-1

Miraldo, A., Sundh, J., Iwaszkiewicz-Eggebrecht, E., Buczek, M., Goodsell, R., Johansson, H., Fisher, B., Raharinjanahary, D., Rajoelison, E., Ranaivo, C., Randrianandrasana, C., Rafanomezantsoa, J.-J., Manoharan, L., Granqvist, E., Dijk, L. van, Alberg, L., Åhlén, D., Aspebo, M., Åström, S., … Ronquist, F. (2024). Data of the Insect Biome Atlas: a metabarcoding survey of the terrestrial arthropods of Sweden and Madagascar. *BioRxiv*, 2024.10.24.619818. https://doi.org/10.1101/2024.10.24.619818

Mora, C., Tittensor, D. P., Adl, S., Simpson, A. G. B., & Worm, B. (2011). How Many Species Are There on Earth and in the Ocean? *PLOS Biology*, *9*(8), e1001127. https://doi.org/10.1371/JOURNAL.PBIO.1001127

Musvuugwa, T., Dlomu, M. G., & Adebowale, A. (2021). Big Data in Biodiversity Science: A Framework for Engagement. In *Technologies* (Vol. 9, Issue 3). MDPI. https://doi.org/10.3390/technologies9030060

Myers, N., Mittermeler, R. A., Mittermeler, C. G., Da Fonseca, G. A. B., & Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature 2000 403:6772*, *403*(6772), 853–858. https://doi.org/10.1038/35002501

Nature Tech Collective. (2024). *Nature Fintech Sector Map*. https://twitter.com/naturex_ntc

Newbold, T., Hudson, L. N., Hill, S. L. L., Contu, S., Lysenko, I., Senior, R. A., Börger, L., Bennett, D. J., Choimes, A., Collen, B., Day, J., De Palma, A., Díaz, S., Echeverria-Londoño, S., Edgar, M. J., Feldman, A., Garon, M., Harrison, M. L. K., Alhusseini, T., … Purvis, A. (2015). Global effects of land use on local terrestrial biodiversity. *Nature 2015 520:7545*, *520*(7545), 45–50. https://doi.org/10.1038/nature14324

Nyström, J. (2024). *Large-scale biodiversity intactness estimation using Bayesian hierarchical models* [Uppsala University]. https://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-538994

Panwar, R., Ober, H., & Pinkse, J. (2023). The uncomfortable relationship between business and biodiversity: Advancing research on business strategies for biodiversity protection. *Business Strategy and the Environment*, *32*(5), 2554–2566. https://doi.org/10.1002/BSE.3139

Pereira, H. M., Ferrier, S., Walters, M., Geller, G. N., Jongman, R. H. G., Scholes, R. J., Bruford, M. W., Brummitt, N., Butchart, S. H. M., Cardoso, A. C., Coops, N. C., Dulloo, E., Faith, D. P., Freyhof, J., Gregory, R. D., Heip, C., Höft, R., Hurtt, G., Jetz, W., … Wegmann, M. (2013). Essential biodiversity variables. *Science*, *339*(6117), 277–278.

Phillips, H., De Palma, A., Gonzalez, R. E., & Contu, S. (2021). The Biodiversity Intactness Index - country, region and global-level summaries for the year 1970 to 2050 under various scenarios. *Natural History Museum Data Portal*. https://doi.org/10.5519/HE1EQMG1

Pyle, R. L., Barik, S. K., Christidis, L., Conix, S., Costello, M. J., van Dijk, P. P., Garnett, S. T., Hobern, D., Kirk, P. M., Lien, A. M., Orrell, T. M., Remsen, D., Thomson, S. A., Wambiji, N., Zachos, F. E., Zhang, Z. Q., & Thiele, K. R. (2021). Towards a global list of accepted species V. The devil is in the detail. *Organisms Diversity and Evolution*, *21*(4), 657–675. https://doi.org/10.1007/S13127-021-00504-0/FIGURES/2

Ranger, N., Alvarez, J., Freeman, A., Harwood, T., Obersteiner, M., Paulus, E., Sabuco, J., Svartzman, R., Althouse, J., Gabet, M., Hurst, I., Ladze, I., Millard, S., Sanchez Juanino, P., David Craig, T., McKenzie, E., Goldner, T., Dutt, N., Baker, L., … de Sousa Almeida, I. (2023). *The Green Scorpion: the Macro-Criticality of Nature for Finance Foundations for scenario-based analysis of complex and cascading physical nature-related financial risks Published as Oxford-NGFS Occasional Paper 2 Citation the Task force on biodiversity loss and nature-related risks (Task force Nature) of the Network of Central Banks and Supervisors for Greening the Financial System (NGFS). We express our sincere thanks to*. https://www.ngfs.net/en/the-green-scorpion-macro-criticality-nature-for-finance

Robinson, J. M., Annells, A., Cavagnaro, T. R., Liddicoat, C., Rogers, H., Taylor, A., & Breed, M. F. (2024). Monitoring soil fauna with ecoacoustics. *Proceedings of the Royal Society B*, *291*(2030). https://doi.org/10.1098/RSPB.2024.1595

Roslin, T., & Laine, A.-L. (2022). The changing fauna and flora of Finland – discovering the bigger picture through long-term data. *Memoranda Societatis pro Fauna et Flora Fennica*, *98*(Supplement 2), 40–53. https://journal.fi/msff/article/view/122353

Ronquist, F., Forshage, M., Häggqvist, S., Karlsson, D., Hovmöller, R., Bergsten, J., Holston, K., Britton, T., Abenius, J., Andersson, B., Buhl, P. N., Coulianos, C. C., Fjellberg, A., Gertsson, C. A., Hellqvist, S., Jaschhof, M., Kjærandsen, J., Klopfstein, S., Kobro, S., … Gärdenfors, U. (2020). Completing Linnaeus's inventory of the Swedish insect fauna: Only 5,000 species left? *PLOS ONE*, *15*(3), e0228561. https://doi.org/10.1371/JOURNAL.PONE.0228561

Roslin, T., Somervuo, P., Pentinsaari, M., Hebert, P. D. N., Agda, J., Ahlroth, P., Anttonen, P., Aspi, J., Blagoev, G., Blanco, S., Chan, D., Clayhills, T., deWaard, J., deWaard, S., Elliot, T., Elo, R., Haapala, S., Helve, E., Ilmonen, J., … Mutanen, M. (2022). A molecular-based identification resource for the arthropods of Finland. *Molecular Ecology Resources*, *22*(2), 803–822. https://doi.org/10.1111/1755-0998.13510

SBTN. (2020). *SCIENCE-BASED TARGETS for NATURE Initial Guidance for Business*.

Seddon, N., Chausson, A., Berry, P., Girardin, C. A. J., Smith, A., & Turner, B. (2020). Understanding the value and limits of nature-based solutions to climate change and other global challenges. *Philosophical Transactions of the Royal Society B*, *375*(1794). https://doi.org/10.1098/RSTB.2019.0120

Seidl, A., Cumming, T., Arlaud, M., Crossett, C., & van den Heuvel, O. (2024). Investing in the wealth of nature through biodiversity and ecosystem service finance solutions. *Ecosystem Services*, *66*, 101601. https://doi.org/10.1016/J.ECOSER.2024.101601

Smith, T., Paavola, J., & Holmes, G. (2019). Corporate reporting and conservation realities: Understanding differences in what businesses say and do regarding biodiversity. *Environmental Policy and Governance*, *29*(1), 3–13. https://doi.org/10.1002/EET.1839

Spear, D., van Wilgen, N. J., Rebelo, A. G., & Botha, J. M. (2023). Collating biodiversity occurrence data for conservation. *Frontiers in Ecology and Evolution*, *11*, 1037282. https://doi.org/10.3389/FEVO.2023.1037282/BIBTEX

Steffen, W., Richardson, K., Rockström, J., Cornell, S. E., Fetzer, I., Bennett, E. M., Biggs, R., Carpenter, S. R., De Vries, W., De Wit, C. A., Folke, C., Gerten, D., Heinke, J., Mace, G. M., Persson, L. M., Ramanathan, V., Reyers, B., & Sörlin, S. (2015). Planetary boundaries: Guiding human development on a changing planet. *Science*, *347*(6223). https://doi.org/10.1126/SCIENCE.1259855/SUPPL_FILE/STEFFEN-SM.PDF

Steenweg, R., Hebblewhite, M., Kays, R., Ahumada, J., Fisher, J. T., Burton, C., Townsend, S. E., Carbone, C., Rowcliffe, J. M., Whittington, J., Brodie, J., Royle, J. A., Switalski, A., Clevenger, A. P., Heim, N., & Rich, L. N. (2017). Scaling-up camera traps: monitoring the planet's biodiversity with networks of remote sensors. *Frontiers in Ecology and the Environment*, *15*(1), 26–34. https://doi.org/10.1002/FEE.1448

Swinfield, T., Shrikanth, S., Bull, J. W., Madhavapeddy, A., & zu Ermgassen, S. O. S. E. (2024). Nature-based credit markets at a crossroads. *Nature Sustainability 2024 7:10*, *7*(10), 1217–1220. https://doi.org/10.1038/s41893-024-01403-w

Stork, N. E. (2018). How Many Species of Insects and Other Terrestrial Arthropods Are There on Earth? *Annual Review of Entomology*, *63*(Volume 63, 2018), 31–45. https://doi.org/10.1146/ANNUREV-ENTO-020117-043348/CITE/REFWORKS

Takano, A., Cole, T. C. H., & Konagai, H. (2024). A novel automated label data extraction and data base generation system from herbarium specimen images using OCR and NER. *Scientific Reports*, *14*(1). https://doi.org/10.1038/s41598-023-50179-0Troudet, J., Grandcolas, P., Blin, A., Vignes-Lebbe, R., & Legendre, F. (2017). Taxonomic bias in biodiversity data and societal preferences OPEN. *Nature*. https://doi.org/10.1038/s41598-017-09084-6

Troudet, J., Vignes-Lebbe, R., Grandcolas, P., & Legendre, F. (2018). The Increasing Disconnection of Primary Biodiversity Data from Specimens: How Does It Happen and How to Handle It? *Systematic Biology*, *67*(6), 1110–1119. https://doi.org/10.1093/SYSBIO/SYY044

IPCC (2023) Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, H. Lee and J. Romero (eds.)]. IPCC, Geneva, Switzerland, pp. 35-115, doi: 10.59327/IPCC/AR6-9789291691647.

TNFD. (2024). *A roadmap for upgrading market access to decision-useful nature-related data*. https://tnfd.global/wp-content/uploads/2024/10/Discussion-paper_Roadmap-for-enhancing-market-access-to-nature-data.pdf?v=1730281144

TNFD. (2023). *Findings of a high-level scoping study exploring the case for a global nature-related public data facility*. https://wedocs.unep.org/bitstream/handle/20.500.11822/41335/state_

TNFD. (2024, December 10). *Over 500 organisations and $17.7 trillion AUM now committed to TNFD-aligned risk management and corporate reporting*. Https://Tnfd.Global/over-500-Organisations-and-17-7-Trillion-Aum-Now-Committed-to-Tnfd-Aligned-Risk-Management-and-Corporate-Reporting/. https://tnfd.global/over-500-organisations-and-17-7-trillion-aum-now-committed-to-tnfd-aligned-risk-management-and-corporate-reporting/

Troudet, J., Grandcolas, P., Blin, A., Vignes-Lebbe, R., & Legendre, F. (2017). Taxonomic bias in biodiversity data and societal preferences. *Nature*. https://doi.org/10.1038/s41598-017-09084-6.

Van Klink, R., August, T., Bas, Y., Bodesheim, P., Bonn, A., Fossøy, F., Høye, T. T., Jongejans, E., Menz, M. H. M., Miraldo, A., Roslin, T., Roy, H. E., Ruczyński, I., Schigel, D., Schäffler, L., Sheard, J. K., Svenningsen, C., Tschan, G. F., Wäldchen, J., … Bowler, D. E. (2022). Emerging technologies revolutionise insect ecology and monitoring. *Trends in Ecology & Evolution*, *37*(10), 872–885. https://doi.org/10.1016/J.TREE.2022.06.001

Van Klink, R., Sheard, J. K., Høye, T. T., Roslin, T., Do Nascimento, L. A., & Bauer, S. (2024). Towards a toolkit for global insect biodiversity monitoring. In *Philosophical Transactions of the Royal Society B: Biological Sciences* (Vol. 379, Issue 1904). Royal Society Publishing. https://doi.org/10.1098/rstb.2023.0101

WEF. (2024). *World Economic Forum Global Risks Report*. www.weforum.org

White, T. B., Mukherjee, N., Petrovan, S. O., & Sutherland, W. J. (2023). Identifying opportunities to deliver effective and efficient outcomes from business-biodiversity action. *Environmental Science & Policy*, *140*, 221–231. https://doi.org/10.1016/J.ENVSCI.2022.12.003

White, T. B., Petrovan, S. O., Bennun, L. A., Butterworth, T., Christie, A. P., Downey, H., Hunter, S. B., Jobson, B. R., zu Ermgassen, S. O. S. E., & Sutherland, W. J. (2023). Principles for using evidence to improve biodiversity impact mitigation by business. *Business Strategy and the Environment*, *32*(7), 4719–4733. https://doi.org/10.1002/BSE.3389

Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J. W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., … Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data 2016 3:1*, *3*(1), 1–9. https://doi.org/10.1038/sdata.2016.18

Wieczorek, J., Bloom, D., Guralnick, R., Blum, S., Döring, M., Giovanni, R., Robertson, T., & Vieglais, D. (2012). Darwin Core: An Evolving Community-Developed Biodiversity Data Standard. *PLOS ONE*, *7*(1), e29715. https://doi.org/10.1371/JOURNAL.PONE.0029715

Whitley, B. S., Abermann, J., Alsos, I. G., Biersma, E. M., Gårdman, V., Høye, T. T., Jones, L., Khelidj, N. M., Li, Z., Losapio, G., Pape, T., Raundrup, K., Schmitz, P., Silva, T., Wirta, H., Roslin, T., Ahlstrand, N. I., & Vere, N. de. (2024). Harmonising digitised herbarium data to enhance biodiversity knowledge: creating an updated checklist for the flora of Greenland. *BioRxiv*, 2024.12.01.626242. https://doi.org/10.1101/2024.12.01.626242

WWF. (2024). *WWF Position Voluntary Biodiversity Credits*. https://wwfint.awsassets.panda.org/downloads/biodiversity-credits-position---october-2024---final.pdf

Supporting Information

Case study S1. Biodiversity Data for Financial Metrics: MISTRA FinBio, Sweden

Figure S1. Photo of Malaise traps in pilot fields, taken by E Granqvist May 2024.

The “Biodiversity Data for Financial Metrics” work package connects biodiversity data to financial decision-making in the FinBio research program hosted at the Stockholm Resilience Centre (FinBio, 2023). The overall program aims to support financial institutions in contributing to biodiversity and nature-positive outcomes. FinBio operates as a collaborative partnership between academic and financial institutions. The program brings together academic and impact partners to develop practical tools that guide investment decisions, promoting both the greening of finance and the financing of green initiatives that can be adopted throughout the financial sector. The program explores several key areas, and the “Biodiversity Data for Financial Metrics” work package focuses on the application of modern monitoring technologies such as environmental DNA (eDNA) and Earth Observations for assessing biodiversity impact. This includes methods for assessing Essential Biodiversity Variables (EBVs). The methodology encompasses eDNA collection from Malaise traps and soil samples, focusing on laboratory and bioinformatic protocols, accuracy measurements, and abundance estimation, with trend analyses covering a five-year period in Sweden. A pilot project within the work package involves collaboration with Svensk Kolinlagring, a non-profit organization launched in 2019 that connects stakeholders to increase soil health and carbon storage in Swedish agricultural soils. This organization currently works with approximately 40 farms. The pilot project with FinBio focuses on measuring biodiversity in agricultural farmland targeting carbon sequestration, which the IPCC has identified as one of the most cost-effective and scalable climate action solutions. The pilot project aims to deliver several key outcomes, including biodiversity data from the agricultural sector using eDNA monitoring methods, analysis of biodiversity changes in carbon sequestration management systems, and the development of a biodiversity index for farmers. This index will serve as both a measurement tool and a component of potential business cases to attract investment in sustainable agricultural practices. Open data and open methods are core principles within the Biodiversity Data for Financial Metrics work package, and the collected pilot data will be shared via GBIF upon completion of the project.

Case study S2. Total Energies share biodiversity data on GBIF

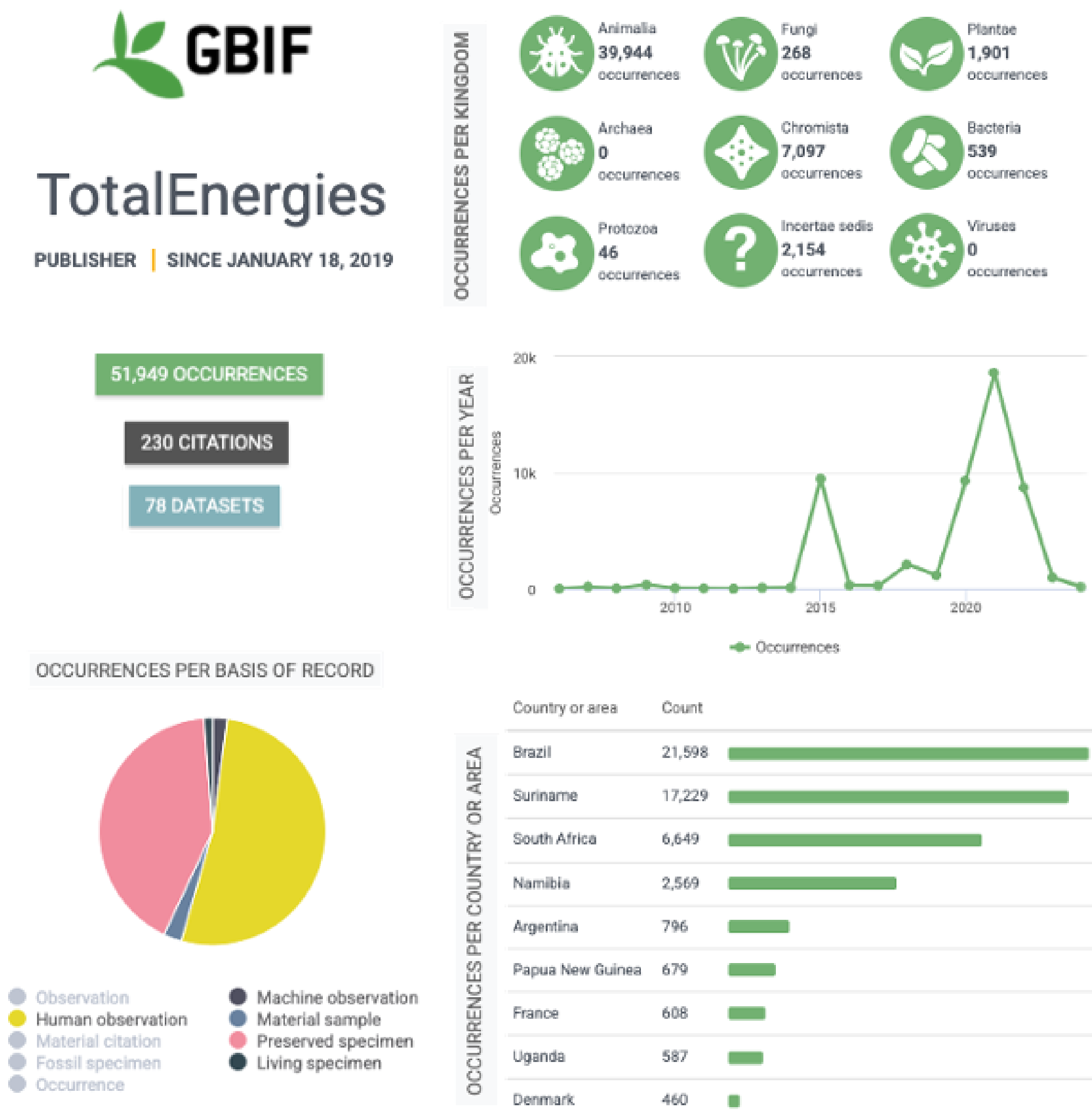

Figure S2. TotalEnergies data publisher metrics displayed on their GBIF publisher page, showcasing key performance indicators (KPIs) for company reporting. Metrics include: occurrences per kingdom, occurrences per year, occurrences per country or area, and occurrences per basis of record. These metrics provide insights into data distribution and can be used to evaluate the company's contribution to biodiversity monitoring.

TotalEnergies, a global energy company operating in 120 countries, became a publisher of the Global Biodiversity Information Facility (GBIF) in 2018 to strengthen its efforts in biodiversity data sharing. The company committed to sharing biodiversity data collected through environmental impact assessments, including field surveys in remote and offshore locations, with both the scientific community and the public. By publishing its data to GBIF, TotalEnergies considers this a valuable contribution to global scientific research and international conservation efforts. The company employs a variety of data collection methods, such as sediment, soil, and water sampling, camera transects, and passive acoustic monitoring and opportunistic observations of marine megafauna and birds. This data encompasses hydrocarbons, metals, microbiology, and benthic fauna, helping to assess habitat sensitivity. TotalEnergies' biodiversity data adheres to GBIF's quality standards, following DarwinCore (DwC) standard and FAIR principles, and has committed to contribute data annually from a minimum of five projects or sites to GBIF, with regular reports on these contributions.

Companies publish biodiversity data by establishing institutional agreements and complying with GBIF's Data Publisher and Data User Agreements. Registration as a data publisher requires endorsement from a national GBIF node. The process typically involves collaboration with contractors and field technicians to ensure data and metadata quality. Companies must establish internal workflows, select and prepare biodiversity data according to the DwC, define access restrictions, and publish under a Creative Commons license.