Capturing the value of biosurveillance "big data" through natural capital accounting

David Castle, Paul D.N. Hebert, Elizabeth L. Clare, Ian D. Hogg & Crystal Tremblay

Published online: 30 Jul 2021.

Submit your article to this journal ☑

Article views: 33

View related articles ☑

View Crossmark data ☑

Capturing the value of biosurveillance "big data" through natural capital accounting

David Castle[a], Paul D.N. Hebert[b], Elizabeth L. Clare[c], Ian D. Hogg[d]
and Crystal Tremblay[e]

[a]School Public Administration and Gustavason School of Business, University of Victoria, Canada; [b]Centre for Biodiversity Genomics, University of Guelph, Guelph, Canada; [c]Ecology and Evolutionary Biology, York University, Toronto, Canada; [d]Ecosystem Science, Polar Knowledge Canada, Ottawa and Cambridge Bay, Canada; [e]Department of Geography, University of Victoria, Canada

ABSTRACT

Global biodiversity is in crises. Recognition of the scale and pace of biodiversity loss is leading to rapid technological development in biodiversity science to identify species, their interactions, and ecosystem dynamics. National and international policy developments to stimulate mitigation and remediation actions are escalating to meet the biodiversity crises. They can take advantage of biosurveillance "big data" as evidence for more sweeping and impactful policy measures. The critical factor is translating biosurveillance data into the value-based frameworks underpinning new policy measures. An approach to this integration process, using natural capital accounting frameworks is developed.

1. Introduction

Scientific and technological capability now exists to provide standardized, internationally comparative, and geospatially linked data on species identification, species interactions, and ecosystem dynamics. Global biodiversity "big data" production can be tapped and used to pursue the goals, targets and indicators of at least three Sustainable Development Goals (SDGs) by 2030, and to support biodiversity and sustainability objectives thereafter. A translation step at the interface between understanding of natural and scientific systems is necessary to develop and implement policy measures. Biodiversity data at scale are critical for answering key scientific questions about species abundance, diversity, and interactions and is the backbone of an emerging global bio-surveillance system, BIOSCAN. Critically, the same data must be interpreted and valued in light of economic, social or cultural considerations for uptake into policy measures. This translational step can be accomplished through systems of natural capital accounting that provide a framework to explore interactions between the natural and social systems. Motivated by the global anthropogenic biodiversity crisis and the socio-economic stakes of inaction, the urgent goal of these frameworks is to provide integrated natural-social evidence for policy measures to protect planetary biodiversity. New emerging technologies, supported

by artificial intelligence and machine learning techniques can take these frameworks one step further by creating "digital twins" and avatars of natural-social systems for testing hypotheses related to policy options for protecting planetary biodiversity.

2. Sustainable Development Goals and biodiversity

Planetary biodiversity is under unprecedented threat as ecosystems are being restructured with corresponding devastating biodiversity loss. The first global extinction event in 65 million years will be anthropogenic and without decisive action an eighth of all species will be lost by 2100 (Dirzo et al., 2014). The World Economic Forum's *Global Risks Report 2020* lists biodiversity loss as the fourth most likely risk with the third most serious impact (after climate action failure and weapons of mass destruction) (World Economic Forum, 2020). Responses to this crisis are being issued by the United Nations Convention on Biological Diversity's Global Biodiversity Outlook 5 (Convention on Biological Diversity, 2020a) and in the Horizon Europe Candidate Partnership on Biodiversity which calls for the investment of €500 M to halt the anthropogenic-driven global decline of species.

Dire anticipated consequences stem from the past failure to fully recognize humanity's dependency on the natural environment. The Millennium Ecosystem Assessment (2005), which began its work more than two decades ago, emphasizes that "people are integral parts of ecosystems and that a dynamic interaction exists between them and other parts of ecosystems with the changing human condition driving, both directly and indirectly, changes in ecosystems and thereby causing changes in human well-being." More recently, the Global Biodiversity Outlook 5 (Secretariat of the Convention on Biological Diversity, 2020) outlines eight interdependent transitions (land and forests; freshwater; fisheries and oceans; sustainable agriculture; food systems; cities and infrastructure; climate action; one health) that will move society towards a more sustainable relationship with nature. Like the Millennium Assessment, each transition area is premised on the human dependence on biodiversity and the consequences of harm to the environment.

Sustainable human dependency on the environment is the fundamental underpinning of the Sustainable Development Goals (SDGs), adopted by all United Nations Member states in 2015. The goals most directly focused on biodiversity are *Goal 14: Conserve and sustainably use the oceans, seas and marine resources for sustainable development* and G*oal 15: Protect, restore and promote sustainable use of terrestrial ecosystems, sustainably manage forests, combat desertification, and halt and reverse land degradation and halt biodiversity loss*. Between these two goals, there are 16 primary and six secondary targets covering diverse objectives and indicators. Additionally, the status of global biodiversity is implicated in *Goal 13: Take urgent action to combat climate change and its impacts*, since biodiversity and ecological impacts, and their redress through mitigation measures, are relevant indicators. The SDGs can only be fully realized with strong partnerships at all levels (SDG 17): Strengthen the means of implementation and revitalize the global partnership for sustainable development, that build upon shared values and inclusion.

Among the SDG Goal 15 targets, 15.9 is focused on biodiversity as it states: "By 2020, integrate ecosystem and biodiversity values into national and local planning, development processes, poverty reduction strategies and accounts." The sub-target 15.9.1 further defines the objective as "Progress towards national targets established in accordance with Aichi Biodiversity Target 2 of the Strategic Plan for Biodiversity 2011–2020." The Aichi

targets were set at the tenth meeting of the Conference of the Parties to the Convention on Biological Diversity in Nagoya in 2010 and were subsequently updated into national biodiversity strategies and action plans (Convention on Biological Diversity, 2021). Target 2 states: "By 2020, at the latest, biodiversity values have been integrated into national and local development and poverty reduction strategies and planning processes and are being incorporated into national accounting, as appropriate, and reporting systems (Convention on Biological Diversity, 2020b).

The 2020 Secretary-General's report on the *Progress Towards the Sustainable Development Goals* (United Nations, 2020a) assessed progress toward national targets related to Aichi Biodiversity Target 2, noting "[a]bout half the parties had made progress towards their targets, but not at a rate that will allow them to meet their goals." The full *Sustainable Development Goals Report* (United Nations, 2020b) reports that wildlife crime, land and forest degradation, and weak protection measures caused two-thirds of countries to fail to meet their targets. Global species extinction worsened by 10% over the last three decades, and the Red List index has declined from 0.82 in 1990 to 0.73 in 2020, and is projected to fall to 0.70 or less by 2030. Progress is also being hampered by data quality issues, primarily inconsistencies in digitization versus paper records, standardization allowing for international comparison, and the "integrated analysis and visualization of geospatially enabled data."

Those nations that are partially on track to meet biodiversity targets have "national and local development plans and national accounting and reporting systems have integrated the values reflected in international treaties and strategic plans on biodiversity (United Nations , 2020b, p. 55)." Data associated with targets and initiatives are a crucial input into government processes that yield these kinds of policy measures. While acknowledging the ongoing need for improved amounts and quality of data to characterize each SDG, and to monitor key indicators of progress (or lack thereof), the policy challenge is to identify and incorporate appropriate indicator data in an evolving national and international policy environment. A long-standing problem, the development of effective policy measures critically depends on improved understandings of how natural and social systems interact. Without a deeper understanding of these interactions, critical gaps in knowledge will prevent policy-relevant questions from being answered and policy measures being developed and implemented.

3. Data – abundance and value

Many research programs generate abundant data. For example, the 3.5 billion base-pair human genome is ~700 MB of data and will fit on a CD-ROM, which scales up quickly when undertaking an initiative like the UK's 100,000 Genomes Project. Since March 2020, when the COVID-19 pandemic was declared, 300,000 papers and related datasets on this virus have been published (CORD-19, 2021). The Hubble Telescope produces about 10 TB of data each year (NASA, 2021), while NASA's earth observing data output approaches 33 TB per day, contributing to an accumulated 50PB of data in the Earth Observing System Data and Information System (NASA, 2020). CERN's data centre produces an average of 1 PB per day with experiments on the Large Hadron Collider generating 90 PB yearly, and non-LHC experiments another 25 PB annually (CERN, 2020). The Square Kilometre Array, the world's largest radio-telescope, generates 300 PB from its thousands of telescopes at both its Australian and South African Nodes) resulting in an output of 8.5 Exabytes over the 15 year planned lifespan for first-round high-priority science programmes (Scaife et al.,

2020). Considering the global research and government science networks, hundreds of exabytes of research data are generated and shared yearly.

In 1997, the concept of "big data" was first framed as a challenge for visualization: " . . . data sets are generally quite large, taxing the capacities of main memory, local disk, and even remote disk. We call this the problem of *big data*" (Cox & Ellsworth, 1997). Big data were a problem because data was scaling in a nonlinear fashion, outstripping computational growth, and presenting qualitatively different computational challenges. The problem has only grown with additional complexities introduced by the surge in inconsistent metadata and computational requirements requiring advents in machine learning and artificial intelligence. Since data can be amassed at scale for many research questions, the contemporary challenge is to corral relevant data and interpret it to provide on-point answers to scientific questions. Data science has arisen as a result, and with it has come the growing need for digital skills and workforce capacity development (OECD, 2020; EC DG-RTD, 2021). The more daunting qualitative challenge is that the " . . . epistemic power of big data lies in their capacity to bridge between different research communities, methodological approaches and theoretical frameworks that are difficult to link due to conceptual fragmentation, social barriers and technical difficulties" (Leonelli, 2020). Not only can big data play a central role in the generation of knowledge and the structure of systems of knowledge, but working across domains of knowledge creates opportunities to tackle problems with scale and complexity for which big data are essential.

Although there are theoretical accounts of observational and causal inference and the nature of evidence (Achinstein, 2001), these are beyond the scope of the present discussion, the purpose of which is to present a pragmatic pathway to use biodiversity data to create evidence for the policy measures needed to prevent further species loss and destabilized ecosystems. Biodiversity science provides deepening insights about the scale and severity of threats to species in which habitat loss is the main cause. Prevention and reversal of habitat loss are obvious objectives, but biodiversity science does not itself chart a course. Instead, sense must be made by translating results from biodiversity science into terms that stand as evidence for policies and decisions that achieve those goals. This is not a new challenge. Assessing the interactions between social and natural systems has proven difficult because scientific and assessment tools needed further development, and because the study of ecological and social systems focus on respective characteristics of each system. The need to develop "all-important interactions between these systems" was pointed out in the Millennium Assessment, but the point can be refined: The issue is to understand how the big data amassed in biodiversity science can be used by other conceptual frameworks to gain "epistemic power" by adding new meaning, value, and salience to that data in ways that lead to policies and decisions that stem biodiversity loss.

4. BIOSCAN and big data for global biosurveillance

BIOSCAN is a global biodiversity genomics program that that was launched in 2019 by the International Barcode of Life Consortium (iBOL), an effort expected to achieve its core goals by 2045. Led by the Centre for Biodiversity Genomics at the University of Guelph, Canada, BIOSCAN builds on nearly two decades of work in developing and applying DNA-based identification systems based on targeted sequencing of short, standardized gene regions (DNA barcodes) (Hebert et al., 2003). Over the years, DNA barcoding has

demonstrated its effectiveness for species identification and disambiguation (Hebert et al., 2016), activities supported by a workbench and online reference library Barcode of Life Data System (BOLD) with more than 9 million barcode records representing more than 800,000 species, and 30,000 species identifications per week (Centre for Biodiversity Genomics, 2021). Core to BOLD is the Barcode Index Number (BIN) system that uses algorithms to support specimen identification and species discovery (Ratnasingham & Hebert, 2007; Ratnasingham, Hebert, & Fontaneto, 2013).

More than 700,000 BIN pages now reside on BOLD, each assembling information on specimens assigned to a particular taxon. Because studies have shown their strong congruence with species recognized through morphological study (Pentinsaari et al., 2019), BINs are an effective species proxy in groups that have received little taxonomic attention. The capacity of the BIN system to automate species recognition is much needed because 80% of multicellular species await description (Mora, Tittensor, Adl, Simpson, & Worm, 2011). Furthermore, completion of this task could require 240 billion USD and 600 years if pursued using morphological approaches (Carbayo & Marques, 2011). By contrast, the BIN system makes it feasible to register all animal species in two decades for less than 1 USD billion, and to extend coverage to include plants and fungi for a modest additional investment.

Existing barcoding platforms demonstrate that DNA sequencing technologies have revolutionized biodiversity science. Given the urgency of understanding the diversity of life because of the threat of mass extinction, the priority now is to scale-up species identification with a workflow that takes advantage of high-throughput sequencers (HTS), normally used for whole genomes, for DNA barcoding. With the shift to HTS platforms, data abundance is not a constraint – short-read (<500 bp) platforms now generate billions of sequences in a run, while their long-read (>10 kb) counterparts produce millions. Workflow development is currently underway to optimize the "big data" output (database, image, DNA extraction, PCR, sequence acquisition/upload/validation, taxonomic/BIN assignment, voucher/DNA curation). A further goal is to reduce costs to 1 USD per specimen by 2023 while improving data quality and increasing production. With this infrastructure, 10 million single specimens will be processed, species interactions will be characterized by examining the symbiomes of another million specimens, while baseline data on species distributions will be developed by analyzing 100,000 bulk samples, containing more than a hundred million specimens from 2000 locations.

New BIOSCAN HTS protocols deliver volume and precision while minimizing costs. These are design points for biosurveillance construed deliberately as a "big data" initiative to accelerate species identification, enable low cost and accurate analysis of species interactions, and develop DNA metabarcoding protocols to assess shifts in biological communities at all scales to address more complex biodiversity metrics. Global biodiversity surveillance – "biosurveillance" – is gaining momentum with 40 nations joining BIOSCAN with confirmed and pending funding exceeding 100 USD M. These nations will contribute variously to upstream R&D to optimize protocols which will be used to create DNA barcode reference libraries. All national efforts have project partners and users of biodiversity knowledge, including the Convention on Biological Diversity (CBD), global conservation NGOs, domestic regulators, and environmental consultants, communities, Indigenous peoples, and the public. BIOSCAN's biosurveillance program has the potential

to inform mitigation strategies to prevent biodiversity loss and ecosystem perturbation sought under the SDGs.

4.1. Transforming biosurveillance

BIOSCAN goes far beyond counting species. This approach is truly transformative illuminating patterns of interactions among species and to tracking the shifting distributions of species in response to environmental change at previously impossible scales. The promise of real-time global reporting enabling real-time crisis intervention and economic accounting will revolutionize our valuations of biodiversity. A rapidly growing economic market for sustainable ecosystems is emerging. For example, issuance of green bonds exceeded all 2019 expectations (Financial Times, 2020) with Moody's revising projections to 250 USDbn. The stage is set for green asset development as part of an international market in natural capital. When combined with technological advances in high throuput sequencing and integrated with digital technologies such as continental spatial mapping (e.g. LIDAR), a truly digital environment for the assessment and subsequent valuation of natural assets will emerge.

The transformation of regulatory processes to adopt these big biodiversity data is illustrated in the United Kingdom by the adoption of recommendations from the UK Natural Capital committee (United Kingdom Natural Capital Committee, 2021) linked to the 25 year plan for natural asset-based assessment. Intergovernmental agreements to integrate sustainability into financial policy have been established in the EU (European Union, 2021) and increasing acceptance of carbon cap and trade systems promise to increasingly marketize natural capital.

It is fundamentally important to have in place a system to use BIOSCAN data as it rapidly becomes available to natural capital accounting so that indices accurately capture changes on the frontlines of biodiversity surveillance. This is particularly important as evidence mounts that functionally intact ecosystems are not always co-extensive with existing protected areas (Plumptre et al., 2021), and protected areas are not immune to transformational change (e.g. Google Earth Timelapse). Already BIOSCANs genetic approach has made it possible to rapidly diagnose metacommunity structure (Bush et al., 2020) and ecological interactions (Pompanon et al., 2012) transforming how we measure everything from ecosystem process (Evans et al., 2016) to managing populations (Pont et al., 2019) and becoming fundamentally integrated into international frameworks for resource management (e.g. https://dnaqua.net/) as part of "next generation biomonitoring" (Makiola et al., 2020).

5. Capturing the value of biosurveillance "big data"

To borrow a metaphor from chemistry, BIOSCAN has reached an activation energy threshold; biosurveillance is digitally enabled, scalable biosurveillance, supported by ML and AI, produces petabytes of data, and reveals profound new knowledge about planetary biodiversity. Yet the ability to use this emerging information in integrated natural-social frameworks for policy measures development remains limited. The challenge is to find the threshold energy at the intersection of natural and social scientific frameworks to make an activated complex that directly serves the SDGs.

The Science Academies of the Group of Seven (G7) nations recently asserted that "[h]umans emerged within the biosphere and are both inseparable from it and fully dependent on it" and [a]lmost every pressing issue for humanity is inextricably linked to biodiversity (Science Academies of the Group of Seven, 2021a)." Among several calls to action, the Academies call for:

> "New approaches to valuing and accounting for biodiversity are required so that economies no longer decouple economic growth from the long-term sustainability of the biosphere. These might include natural capital accounting, green investments, ecosystem service valuation, nature-related financial disclosures and other forms of national and corporate accounting that change the behaviours of companies and investors (Science Academies of the Group of Seven, 2021a)."

The first of their three recommendations is to adopt "new approaches to valuing and accounting for biodiversity." The intent is to recognize the multiple ways biodiversity is valued by different communities and cultures around the world. Going further and building on these valuations, the idea of an accounting framework is to articulate how the value of biodiversity underpin conceptions of wellbeing, how decisions are made, inequalities reduced, and how the costs of protecting biodiversity can be internalized in economies. With respect to the latter point, new approaches to accounting for biodiversity are inclusive of the central tenets of circular economies and strive ensure that "biodiversity is addressed in national and corporate accounting procedures . . .". These new procedures are intended to do the work that measures of national wealth through Gross Domestic Product (GDP) which could never do for the SDGs (Costanza et al., 2014).

In his recent, expansive review, *The Economics of Biodiversity*, Partha Dasgupta encourages the economics of biodiversity as a study of portfolio management in which we are all natural asset managers almost all of the time (Dasgupta, 2021). This approach to biodiversity economics is meant more literally than metaphorically. The intent is to become systematic in portfolio management as a methodology, using Natural Capital Accounting (NCA) frameworks as the main method. NCAs are structured approaches that build upon ecosystem goods and services (EGS) measurement techniques. NCAs take the four main service types (provisioning, regulation, cultural supporting) and assimilating the measured value of numerous and often disparate factors. This can be undertaken using a scheme for accounting for ecosystem services, such as the European Environment Agency's Common International Classification of Ecosystem Services (European Environment Agency, 2021) and accounted for in the UN System of Environmental Economic Accounting (SEEA) (United Nations, 2021). The SEEA system has a central framework based on environmental flows, stocks of environmental assets, and economic activity related to the environment. It provides standards for ecosystem accounting in terms of assets, their condition, services, benefits, and considers the beneficiaries such a people, households or businesses.

In general, an NCA creates an "environmental profit and loss account" resembling an income statement, and balance sheets showing the gains and losses of natural assets viewed from both their quality and quantity. The recent Bateman and Mace's NCA framework (2020) is distinguished by the efficient accounting for alternative decisions, by its ability to quickly compare the relative sustainability of options, and by its capacity to track distributional outcomes of decisions to evaluate their equity. It is feasible to

incorporate BIOSCAN outputs into this NCA framework, but it will take significant effort. Developing an NCA framework that optimizes biosurveillance data is within reach in a few years, but to ensure uptake in SDG processes, it is vitally important to consider how socio-economic factors will be considered as policy measures are developed and implemented. Although many factors need consideration, four have primacy: 1) local and global reach with communities, 2) access to internationally comparative data, 3) ensuring uptake into policy processes, and 4) preemptive testing of scenarios using digital twins and avatars.

5.1. Community engagement

The seventh World Report on Higher Education (Global University Network for Innovation (GUNi), 2020) makes a compelling and essential case for coupling the humanities with science and technological innovations as both drivers and critics towards social and planetary transformations. In addition to calling for a new paradigm in the relation between science, technology, and the humanities, the report argues for a multiplicity of knowledges beyond the current Western paradigm, proposing new ways of merging knowledge systems to address pressing ecological and societal challenges. Local communities are best situated to provide direct insights about their attributions of value to biodiversity and what biodiversity shifts and losses mean to them.

In adopting this perspective, the Intergovernmental Platform on Biodiversity and Ecosystem Services (IPBES) expanded upon the ecosystem services concept from the Millennium Assessment to develop a perspective on "nature's contribution to people (NCP)" in which the role of culture in defining links to nature is recognized, an action which "elevates, emphasizes and operationalizes the role of indigenous and local knowledge (Diaz et al., 2018)." The contribution of biodiversity science to our understanding of the stocks and flows of ecosystem services provides a critical foundation of knowledge. This knowledge gains potency when it is interpreted in light of indigenous and local needs, perspectives, histories, and cultural associations with the natural world. Consequently, the relevance and impact of an NCA framework depends on first taking a systematic approach to working with affected communities, taking consideration of different knowledge sources and perspectives, creating opportunities for the co-production of knowledge and cross-cultural valuation work that links people to biodiversity in place.

How does this work in practice? Engagement of indigenous people in the development of an inclusive NCA is paramount, as they have the deepest historical and cultural connections to the natural environment. Indigenous peoples' historical connection to their territories and culturally significant sites and species, and the expectations and obligations in the UN Declaration on the Rights of Indigenous Peoples (United Nations, 2007) must be recognized and respected. Among the many approaches to indigenous and local community engagement, the community-based and indigenous-led research based on protocols developed through the UNESCO Knowledge for Change (K4C) Consortium has gained leadership. This participatory research process emphasizes knowledge co-creation, reciprocity, and horizontal decision-making (Gilmore & Young, 2012; Hall & Tandon, 2020; Mulrennan, Rodney, & Scott, 2012), resulting in shared ownership of the research process with emphasis on community action and knowledge bridging. Indigenous peoples have been at the forefront of protection of ancestral land and waters throughout the world and several recent studies are demonstrating the rich and local-

knowledge systems of Indigenous-led stewardship and conservation practices (Mantyka-Pringle et al., 2017). These studies show strong empirical insights into how knowledge bridging, or "two-eyed seeing" can contribute to more adaptive and locally contextualized co-management practices and solutions for more sustainable and resilient ecosystems (Reid et al., 2020).

Furthermore, on the question of who has access to biodiversity data that is being generated at unprecedented scale and speed, the proponents of the K4C Consortium have argued for a conception of "openness" that supersedes conventional and partial accounts of open science in an effort to democratize and decolonize knowledge (Chan, Hall, Piron, Tandon, & Williams, 2020). The practical implication that follows is that biodiversity data for NCAs cannot be proprietary or closed, and that proactive efforts must be made to share knowledge with communities that contributed to the knowledge and will be affected by biodiversity change and loss.

5.2. Access to internationally comparative data

A major obstacle to the SDGs is the lack of internationally comparative data, but even if there were such data, the foregoing discussion about democratizing openness emphasizes the criticality of data access. As the largest biosurveillance program ever undertaken, BIOSCAN is poised to have a profound global impact through open data access. The short sequences of DNA used in barcoding are compliant with international access and benefits sharing (ABS) protocols construed under the Nagoya Protocol to the Convention on Biological Diversity (CBD), making barcoding less controversial than whole-genome sequencing. Nevertheless, the 2019 First Global Dialogue on Digital Sequence Information on Genetic Resources, identified emergent disputes which could provoke restricted access to digital sequence information (DSI). Parties to the CBD from biodiversity-rich nations propose tight restrictions on access to DSI. Many parties to the CBD from the global south and north oppose this plan, so the debate continues.

Clarity on ABS and DSI related to DNA barcoding is critical to safeguard the norms (e.g. as embodied in the FAIR, TRUST and CARE frameworks) of global open science, as embodied by the soon-to-be finalized UNESCO Recommendation on Open Science (UNESCO, 2021), without which the impact of biodiversity's "big data" and biosurveillance technology will not achieve their full potential for meeting SDG biodiversity targets and goals. Previous work arising from the International Barcode of Life program (Schindel et al., 2015) shows that a critical factor in ABS negotiations is to engage with communities and policymakers early to anticipate ABS challenges before they become intractable problems. This approach to the Nagoya Protocol has limited outright biopiracy and clarified the implementation of ABS with respect to genetic resources, and can apply to anticipate and address DSI considerations.

How does this work in practice? An example for Canada may be illustrative of the need to resolve upstream inconsistencies in practices and policies to promote appropriate access to biosurveillance data. Canada has long supported the CBD but is not a signatory to the Nagoya Protocol. Meanwhile, the federal government has adopted a Nation-to-Nation approach to Canada and First Nations relations. With nation-status there has been an assertion of rights to indigenous data sovereignty, for which the First Nations Information Governance Centre (FNIGC) has created the First Nations principles of OCAP (ownership, control, access and possession). Although less formalized than OCAP,

similar principles apply for partnership in research with the Inuit, described in the National Inuit Strategy on Research, "priority area 4." Métis Nation organizations do not have a similar research guidance, but it is common practice to engage with the Métis using other established protocols, some university-based, and reference to OCAP. Working with communities to develop a model DSI protocol for the BIOSCAN NCA instantiates the community engagement described above, and will ultimately inform ABS/DSI policy development in Canada, an advance that may serve as a model elsewhere.

5.3. Ensuring uptake into policy processes

By design, NCAs have policy relevance with anticipated uses in natural capital appraisals, informing decisions, and monitoring the implementation of decisions. Biosurveillance has enormous policy relevance for science-based departments and agencies that have statutory obligations as regulators (e.g. environmental impact assessments) or which use biodiversity science in other domestic and international policy processes. In an era where policy decisions and their impacts are increasingly complex, good governance depends on policy processes that are initiated with the expectation that natural and social scientific insights and objectives will converge (Lyall, 2005). Because this cannot be left to unfold naturally, there must be proactive efforts to identify the correct government authority (Doern, Castle, & Phillips, 2016; Phillips & Castle, 2021) and to engage strategically across multiple levels of government (Oliver and Cairney, 2019). NCAs are not widely adopted in policy practices worldwide, suggesting that policy relevance will require iterative work with policymakers, particularly since as Bateman and Mace (2020) acknowledge, when the framework is used by decision makers, "the definition of 'relevant effects' may differ." As with all policy processes, success in developing measures depends on having well-structured policy problems from the outset (i.e. general agreement about facts, values, processes, and objectives) (Hoppe, 2010) and constructive approaches to resolve dissent when it arises (Castle & Culver, 2013).

How does this work in practice? One illustrative example arises from HTS DNA barcoding and post-processing symbiome analysis of plant-pollinator interactions, many of which are vitally important in agriculture and contribute significantly to crop value. The broader focus on pollinators includes their role in supporting plant populations beyond agricultural systems, as well as the fungal and viral pathogens. The IPBES comments on the significance plant-pollinator interactions have for local and indigenous communities who regard themselves as having a duty of care to sustain the environments supportive of plant-pollinator interactions, and who hold some species as "totemic", requiring special relationships with people (Intergovernmental Platform on Biodiversity and Ecosystem Services (IPBES), 2019). The policy relevance of biosurveillance of these symbiomes is increased when different ways of valuing these species and their interactions are factored into an NCA because it provides credible, contextualized, and documented inputs in the NCA. Legitimized through lived experience that may include cross-cultural learnings, the biosurveillance knowledge in an NCA is much more likely to provide the foundation for policy measures that will be perceived as meaningful and relevant to engaged communities. In this respect, biosurveillance data from BIOSCAN plays a key role in understanding the species and dynamics of pollinator systems and their potential for impact on human dependence on pollinators using an NCA that is based on the SEEA framework and the Bateman and Mace approach to policy relevance.

5.4. Digital twins and avatars

Policy measures that have impact on SDGs targets for biodiversity will have to conquer major challenges. For example, the Science Academies of the G7 nations observes that while agriculture, forestry, and land use account for 25% of CO_2 emissions, that same conversion of habitat is responsible for most biodiversity loss (Science Academies of the Group of Seven, 2021b). The required "thoughtful action" includes "the sustainable intensification of agriculture, improving soil management to ensure carbon uptake, and making changes to our diet." This is no simple challenge, and the question is whether all three outcomes can be simultaneously achieved given that zero-sum or net-negative outcomes thwart progress towards the SDG goals. Answering this question demands qualitative and quantitative research to jointly inform the definition and development of metrics in the formal accounting structure of the NCA, including the multi-stakeholder informed process of assigning values to the elements of the NCA. Making sense of abundant, diverse data elements will require the extensive use of digital tools as envisioned by the Coalition for Digital Environmental Sustainability (CODES) which launched in March 2021 (United Nations Environment Program (UNEP), 2021). The CODES initiative, which supports the broader objectives of the UN Roadmap for Digital Cooperation (United Nations, 2020c) seeks a reciprocal engagement of environmental responses with digital technologies to provide global environmental intelligence, improve the sustainability of products, services and value chains, use digital technologies to transform sectors, propagate social innovation for sustainability, and control digital rebound effects of more intensive digital technology use.

How does this work in practice? Importantly, there are standardized approaches like the UN SEEA to measure the stocks and flows of ecological goods and services that represent core inputs into an NCA. Proprietary approaches have been developed by the "big four" global consultancies (Deloitte, Ernst & Young, KPMG, PWC) and smaller "boutique" consultancies, such as the UK's Economics for the Environment (eftec). There are even registered International Organization for Standardization (ISO) approaches – 14,007 (Environmental management: Determining environmental costs and benefits) and ISO 14,008 (Monetary valuation of environmental impacts and related environmental aspects). Second, from standardized inputs and an established framework like that of Bateman and Mace, an NCA is a model system. Efforts to model human wellbeing in light of the SDGs in an integrative and testable manner has been undertaken (Costanza et al., 2016; Victor, 2019). More recent visualization techniques, similar to the European Union's effort to develop "Destination Earth," (European Commission Directorate-General for Research and Innovation, 2021), are creating models systems that are "digital twins" to map climate change and experiment with mitigation strategies. A "digital twin" or "extended reality" visualization can be developed for the NCA. An NCA digital twin would be an emerging application of digital twin technologies and could be used to explore fundamental governance questions about the contribution of BIOSCAN biosurveillance data to modelling policy measures. Evolving standards and technical specifications for creating digital twins make them more capable for handling big data, artificial intelligence and machine learning, but can also emphasize re-use, interoperability, maintainability and extensibility (Moyne et al., 2020). The Geneva Science and Diplomacy Estimator (2020) has suggested that integrated digital ecosystem "avatars" should be developed to model natural-human systems. Avatars

are essentially scenarios for testing hypotheses about environmental change resulting from policy choices. An "NCA visualiztion laboratory" would provide the advanced research computing infrastructure and highly qualified personnel to create digital twins and test highly interconnected system avatars. As GESDA comments, this "transdisciplinary understanding is vital to addressing the grand challenges facing society in the 21st century."

6. Conclusion

The SDG goals, indicators and targets require accessible and internationally compatible data. The ability to generate biodiversity "big data" is revolutionizing biodiversity and ecological sciences, and can provide global user-friendly access to standardized data. Furthermore, global efforts to create a biosurveillance network, BIOSCAN, directly links new technological capacity to addressing the SDGs. For many decades, the account of the contribution of science and technology to meet environmental objectives would conclude with an invocation for better policy uptake. Over the past decade, natural capital accounting techniques have matured sufficiently that they can reliably capture the value of the surge in biodiversity "big data". We also are seeing new methods for addressing differential valuations of biodiversity arising in communities and from indigenous peoples, access and benefits sharing needs, and ensuring policy relevance. The relevance and potential impact of policy measures and their implementation are amenable to modelling and virtual realization techniques associated with smart manufacturing and Industrial Revolution 4.0 – namely digital twinning and the exploration of digital avatars to conduct policy scenario analysis. The immediate future goal is to develop digital twins of natural capital accounting frameworks to generate a comparable transition in the policy uptake of biosurveillance big data. We anticipate that digitally twinned natural accounting frameworks will support predictive modelling of policy measures, the development of heuristics such as avatars that reflect policy implementation and evaluation paradigms, and ultimately speed progress toward the SDGs.

Disclosure statement

Notes on contributor

Dr. David Castle is a Professor in the School of Public Administration and the Gustavson School of Business at the University of Victoria and a Researcher in Residence at the Office of the Chief Science Advisor to the Prime Minister of Canada.

His research is focused on science, technology and innovation policy, with a particular emphasis on regulation, standards, intellectual property and public consultation associated with life science innovation.

References

Achinstein, P. (2001). *The book of evidence*. Oxford: Oxford University Press.

Bateman, I. J., & Mace, G. M. (2020). The natural capital framework for sustainably efficient and equitable decision making. *Nature Sustainability*, *3*(10), 776–783.

Bush, A., Monk, W. A., Compson, Z. G., Peters, D. L., Porter, T. M., Shokralla, S., . . . Baird, D. J. (2020). DNA metabarcoding reveals metacommunity dynamics in a threatened boreal wetland wilderness. *Proceedings of the National Academy of Sciences*, *117*(15), 8539–8545.

Carbayo, F., & Marques, A. C. (2011). The costs of describing the entire animal kingdom. *Trends in Ecology & Evolution*, *26*(4), 154–155.

Castle, D., & Culver, K. (2013). Getting to 'no': The method of contested exchange. *Science & Public Policy*, *40*(1), 34–42.

Centre for Biodiversity Genomics (CBG). (2021). BOLD: Barcode of life data systems. Retrieved from https://biodiversitygenomics.net/projects/bold

CERN. (2020). Storage. Retrieved from https://home.cern/science/computing/storage

Chan, L., Hall, B., Piron, F., Tandon, R., & Williams, L. (2020). Open science beyond open access: For and with communities. Canadian commission for UNESCO. Retrieved from https://en.ccunesco.ca/-/media/Files/Unesco/Resources/2020/07/OpenScienceDecolonizingKnowledge.pdf

Convention on Biological Diversity (CBD). (2020a). Global biodiversity outlook 5 (GBOS5): Summary for policymakers. Montreal: Secretariat for the Convention on Biological Diversity. Retrieved from https://ec.europa.eu/digital-single-market/en/destination-earth-destine .

Convention on Biological Diversity (CBD). (2020b). Aichi biodiversity targets. Retrieved from https://www.cbd.int/sp/targets

Convention on Biological Diversity (CBD). (2021). Latest national biodiversity strategies and action plans (NGSAPS). Retrieved from https://www.cbd.int/nbsap/about/latest

CORD-19. (2021). COVID-19 open research dataset. Retrieved from https://www.semanticscholar.org/cord19

Costanza, R., Daly, L., Fioramonti, L., Giovannini, E., Kubiszewski, I., Mortensen, L. F., . . . Wilkinson, R. (2016). Modelling and measuring sustainable wellbeing in connection with the UN sustainable development goals. *Ecological Economics*, *130*, 350–355.

Costanza, R., Kubiszewski, I., Giovannini, E., Lovins, H., McGlade, J., Pickett, K. E., . . . Wilkinson, R. (2014). Time to leave GDP behind. *Nature*, *505*(7483), 283–285.

Cox, M., & Ellsworth, D. (1997). *Application-controlled demand paging for out-of-core-visualization*. IEEE VIS '97: Proceedings of the 8th conference on Visualization. DOI: 10.1109/VISUAL.1997.663888.

Dasgupta, P. (2021). The economics of biodiversity: The Dasgupta review. Retrieved from www.gov.uk/official-documents .

Díaz, S., Pascual, U., Stenseke, M., Martín-López, B., Watson, R. T., Molnár, Z., . . . Shirayama, Y.; Intergovernmental Platform on Biodiversity and Ecosystem Services (IPBES). (2018). Assessing nature's contributions to people. *Science*, *359*(6373), 270–272.

Dirzo, R., Young, H. S., Galetti, M., Ceballos, G., Isaac, N. J. B., & Collen, B.. (2014). Defaunation in the Anthropocene. *Science*, *345*(6195), 401–406.

Doern, B., Castle, D., & Phillips, P. W. B. (2016). *Canadian science, technology and innovation policy: The innovation economy and society nexus*. Montreal: McGill-Queen's University Press.

European Commission Directorate-General for Research and Innovation. (2021). Digital skills for FAIR and open science. doi:10.2777/59065. Retrieved from: https://op.europa.eu/en/publication-detail/-/publication/af7f7807-6ce1-11eb-aeb5-01aa75ed71a1/language-en

European Environment Agency (EEA). (2021). Common international classification of ecosystem services (CICES). Retrieved from https://cices.eu/resources

European Union. (2021). Sustainable finance. Retrieved from https://ec.europa.eu/info/business-economy-euro/banking-and-finance/sustainable-finance_en

Evans, D. M., Kitson, J. J. N., Lunt, D. H., Straw, N. A., Popock, M. J. O., & Poisot, T. (2016). Merging DNA metabarcoding and ecological network analysis to understand and build resilient terrestrial ecosystems. *Functional Ecology*, *30*(12), 1904–1916.

Financial Times. (2020). Green bonds set to keep flying off shelves in 2020. Retrieved from https://www.ft.com/content/61631c2c-1a65-11ea-9186-7348c2f183af

Geneva Science and Diplomacy Estimator (GESDA). (2020). Scientific Anticipatory Briefs. Retrieved from https://gesda.global/scientific-anticipatory-briefs/ .

Gilmore, M. P., & Young, J. C. (2012). The use of participatory mapping in ethnobiological research, biocultural conservation, and community empowerment: A case study from the Peruvian Amazon. *Journal of Ethnobiology*, *32*(1), 6–29.

Global University Network for Innovation (GUNi). (2020). Humanities and higher education: Synergies between science, technology and humanities. Retrieved from http://www.guninetwork.org/report/higher-education-world-7 .

Hall, B., & Tandon, R. (2020). Knowledge democracy for a transforming world. *International Journal of Community Research and Engagement*, *13*, 1–5.

Hebert, P. D. N., Cywinska, A., Ball Shelley, S. L., & deWaard, J. R. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *270*(1512), 313–321.

Hebert, P. D. N., Ratnasingham, S., Zakharov, E. V., Telfer, A. C., Levesque-Beaudin, V., Milton, M. A., & deWaard, J. R. (2016). Counting animal species with DNA barcodes: Canadian insects. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1702), 20150333.

Hoppe, R. (2010). *The governance of problems: Puzzling, powering and participation*. Bristol, UK: Policy Press.

Intergovernmental Platform on Biodiversity and Ecosystem Services (IPBES). (2019). Summary for policymakers of the global assessment report. Retrieved from https://www.ipbes.net/sites/default/files/downloads/spm_unedited_advance_for_posting_htn.pdf .

Leonelli, S. (2020). *Scientific research and big data. In the Stanford encyclopedia of philosophy* (Summer 2020). E. Zalta ed.. Stanford, Stanford University. Retrieved from: https://plato.stanford.edu/archives/sum2020/entries/science-big-data

Lyall, C. (2005). *New modes of governance: Developing an integrated policy approach to science, technology, risk and the environment*. London: Routledge.

Makiola, A., Compson, Z. G., Baird, D. J., Barnes, M. A., Boerlijst, S. P., Bouchez, A., . . . Bohan, D. A. (2020). Key questions for next-generation biomonitoring. *Frontiers in Environmental Science*, *7*, 197.

Mantyka-Pringle, C. S., Jardine, T. D., Bradford, L., Bharadwaj, L., Kythreotis, A. P., Fresque-Baxter, J., & Lindenschmidt, K.-E.; Mantyka-Pringle. (2017). Bridging science and traditional knowledge to assess cumulative impacts of stressors on ecosystem health. *Environment International*, *102*, 125–137.

Millennium Ecosystem Assessment. (2005). *Ecosystems and human well-being: Synthesis*. Washington, DC: Island Press.

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.

Moyne, J., Balta, E. C., Kovalenko, I., Faris, J., Barton, K., & Tilbury, D. M. (2020). A requirements driven digital twin framework: Specification and opportunities. *IEEE*, *8*, 107781–107801.

Mulrennan, M. E., Rodney, M., & Scott, C. H. (2012). Revamping community-based conservation through participatory research. *The Canadian Geographer*, *56*(2), 243–259.

NASA. (2020). Earth science data systems program highlights 2020. Retrieved from https://cdn.earthdata.nasa.gov/conduit/upload/15949/ESDS_Highlights_2020.pdf .

NASA. (2021). About the hubble space telescope. Retrieved from https://www.nasa.gov/mission_pages/hubble/story/index.html .

OECD. (2020). Building digital workforce capacity and skills for data-intensive science. Retrieved from https://read.oecd.org/10.1787/e08aa3bb-en?format=pdf .

Oliver, K., & Cairney, P. (2019). The dos and don'ts of influencing policy: A systematic review of advice to academics. *Palgrave Communications*, *5*(1), 21.

Pentinsaari, M., Anderson, R., Borowiec, L., Bouchard, P., Brunke, A., Douglas, H., . . . Hebert, P.. (2019). DNA barcodes reveal 63 overlooked species of Canadian beetles (Insecta: Coleoptera). *ZooKeys*, *894*, 53–150.

Phillips, P. W. B., & Castle, D. (2021). *Ideas, institutions, and interests: The drivers of Canadian provincial science, technology, and innovation policy*. Toronto: University of Toronto. Forthcoming under contract.

Plumptre, A. J., Baisero, D., Belote, R. T., Vázquez-Domínguez, E., Faurby, S., Jędrzejewski, W., . . . Boyd, C. (2021). Where might we find ecologically intact communities? *Frontiers in Forests and Global Change*, *4*, 26.

Pompanon, F., Deagle, B. E., Symondson, W. O. C., Brown, D. S., Jarman, S. N., & Taberlet, P. (2012). Who is eating what: Diet assessment using next generation sequencing. *Molecular Ecology*, *21*(8), 1931–1950.

Pont, D., Valentini, A., Rocle, M., Maire, A., Delaigue, O., Jean, P., & Dejean, T. (2019). The future of fish-based ecological assessment of European rivers: From traditional EU water framework directive compliant methods to eDNA metabarcoding-based approaches. *Fish Biology*, *98*(2), 354–366.

Ratnasingham, S., & Hebert, P. D. N. (2007). BOLD: The barcode of life data system. Retrieved from www.barcodinglife.org .

Ratnasingham, S., Hebert, P. D. N., & Fontaneto, D. (2013). A DNA-based registry for all animal species: The Barcode Index Number (BIN) System. *PLoS ONE*, *8*(7), e66213.

Reid, A., Eckert, L. E., Lane, J.-F., Young, N., Hinch, S. G., Darimont, C. T., . . . Marshall, A.. (2020). "Two-Eyed Seeing": An indigenous framework to transform fisheries research and management. *Fish & Fisheries*, *22*(2), 243–261.

Rittel, H., & Webber, M. (1973). Dilemmas in a general theory of planning. *Policy Sciences*, *4*(2), 155–169.

Scaife AMM. (2020) Big telescope big data: towards exascale with the Square Kilometre Array. Phil. Trans. R. Soc. A 378: 20190060. http://dx.doi.org/10.1098/rsta.2019.0060 .

Schindel, D. E., Bubela, T., Rosenthal, J., Castle, D., Du Plessis, P., & Bye, R. P. (2015). The new age of the Nagoya protocol. *Nature Conservation*, *12*, 45–56.

Science Academies of the Group of Seven. (2021a). Reversing biodiversity loss – The case for urgent action. Retrieved from https://royalsociety.org/about-us/international/international-work/g-science-academies-meetings .

Science Academies of the Group of Seven. (2021b). A net zero climate-reslient future – Science, technology and the solutions for change. Retrieved from https://royalsociety.org/about-us/international/international-work/g-science-academies-meetings .

Scaife AMM. 2020 Big telescope, big data: towards exascale with the Square Kilometre Array. Phil. Trans. R. Soc. A 378: 20190060. http://dx.doi.org/10.1098/rsta.2019.0060

Secretariat of the Convention on Biological Diversity (CBD). (2020). Global biodiversity outlook 5. Montreal: Convention on Biological Diversity. Retrievd from https://www.cbd.int/gbo5 .

UNESCO. (2021). UNESCO recommendation on open science. Retrieved from https://en.unesco.org/science-sustainable-future/open-science/recommendation .

United Kingdom Natural Capital Committee. (2021). Natural Capital Committee advice on government's 25 Year Environment Plan and progress reports. Retrieved from https://www.gov.uk/government/publications/natural-capital-committee-advice-on-governments-25-year-environment-plan .

United Nations (UN). (2007). Declaration on the Rights of Indigenous Peoples. Retrieved from: https://www.un.org/development/desa/indigenouspeoples/declaration-on-the-rights-of-indigenous-peoples.html .

United Nations. (2021). System of environmental economic accounting (SEEA). Retrieved from https://seea.un.org/content/homepage .

United Nations (UN). (2020a). *Progress Towards the Sustainable Development Goals: Report of the Secretary-General*. Retrieved from https://sustainabledevelopment.un.org/content/documents/26158Final_SG_SDG_Progress_Report_14052020.pdf .

United Nations (UN). (2020b). The sustainable development goals report. Retrieved from https://unstats.un.org/sdgs/report/2020/ .

United Nations (UN). (2020c). Roadmap for digital cooperation. Retrieved from https://www.un.org/es/content/digital-cooperation-roadmap/assets/pdf/Roadmap_for_Digital_Cooperation_EN.pdf .

United Nations Environment Program (UNEP). (2021). Coalition for digital environmental sustainability (CODES). Retrieved from https://www.unep.org/events/webinar/launch-coalition-digital-environmental-sustainability-codes .

United Nations (UN). (2007) Declaration on the Rights of Indigenous Peoples. Retrieved from: https://www.un.org/development/desa/indigenouspeoples/declaration-on-the-rights-of-indigenous-peoples.html

Victor, P. A. (2019). *Managing without growth: Slower by design, not disaster* (2nd ed.). Cheltenham UK: Edward Elgar.

World Economic Forum. (2020). The Global Risks Report. Retrieved from http://www3.weforum.org/docs/WEF_Global_Risk_Report_2020.pdf