SARGASSO SEA COMMISSION

Governance of High Seas Ecosystems: Big Data & AI

Interim Report | 21st December 2021

Produced by:

NLA INTERNATIONAL

BLUE ECONOMY SOLUTIONS

Government Offices of Sweden
Ministry of the Environment

IUCN

Document Information

Client	Sargasso Sea Commission / IUCN / Swedish Government
Project Title	Governance of High Seas Ecosystems: Big Data & AI
Deliverable Number	D01
Report Title	Governance of High Seas Ecosystems: Big Data & AI
Report Version	1.1 (Post SSC Expert & User Group Review)
Report Version Date	21 December 2021
Lead Authors	Name: Dr Kieran N Bjergstrom & Kevin P Fleming Organisation: NLA International Ltd
Contributing Author	Name: Oliver Stringer-Jakins Organisation: NLA International Ltd
Project Manager	Kevin P Fleming Kevin.Fleming@nlai.blue NLA International Ltd
Circulation	1. Client 2. Project Files
File Name	20211221_Governance_of_High_Seas_Ecosystems-Big_Data_&_AI_Interim_Report-FINAL
File Tags	High Seas Governance; Ocean Governance; Maritime Policy; Big Data; Artificial Intelligence; AI; Remote Sensing; ABNJ; BBNJ; Sargasso Sea; Illegal Fishing; IUU Fishing; Ocean Management; Ecosystem Analysis; Ocean Ecosystem Assessment; Marine Spatial Planning; Data Sharing; Data Standards; Maritime Technologies; Earth Observation; Space-Based Sensors; Vessel Monitoring; Underwater Sensors; Blue Economy; Machine Learning; Deep Learning; Image Recognition; Seabed Mining; Technologization; FAIR Data; SIDS; RFMOs; Sargasso Sea Commission; IUCN; NLAI.
File Location	NLA International Ltd: SharePoint

Table of Contents

Glossary

ABNJ	Areas Beyond National Jurisdiction
AI	Artificial Intelligence
AIS	Automatic Identification System
BBNJ	Biodiversity Beyond National Jurisdiction
BE	Blue Economy
BEKS	Blue Economy Knowledge System
CCAMLR	Convention on the Conservation of Antarctic Marine Living Resources
EDA	Ecosystem Diagnostic Analysis
EEZ	Exclusive Economic Zone
EM	Electromagnetic
FFEM	French Global Environment Fund
GEF	Global Environment Facility
GFW	Global Fishing Watch
ICCAT	International Commission for the Conservation of Atlantic Tunas
ICM	Integrated Coastal Management
IMO	International Maritime Organization
IUCN	International Union for Conservation of Nature
IUU	Illegal, Unreported & Unregulated (Fishing)
MPA	Marine Protected Area
MSP	Marine Spatial Planning
NGO	Non-Governmental Organisation
OSPAR	Oslo Paris Commission
PEMSEA	Partnerships in Environmental Management for the Seas of East Asia
RFMO	Regional Fisheries Management Organisations
SAR	Synthetic Aperture Radar
SDG	Sustainable Development Goal
SIDS	Small Island Developing States
SSC	Sargasso Sea Commission
UN	United Nations
UNCLOS	United Nations Convention on the Law of the Sea
UNDP	United Nations Development Programme
VIIRS	Visible Infrared Imaging Radiometer Suite
VMS	Vessel Monitoring System
WCPFC	Western and Central Pacific Fishing Convention

1 Background

The Sargasso Sea is a 2 million square mile open ocean high seas ecosystem. In 2022, the Sargasso Sea Commission (SSC) will be embarking on a major Ecosystem Diagnostic Analysis (EDA) financed by grants from the Global Environment Facility (GEF) and the French Global Environment Fund (FFEM) with the support of a wide number of partners including the currently ten Government Signatories[1] to the 2014 Hamilton Convention on Collaboration for the Conservation of the Sargasso Sea. In support of this project, the Swedish Government has mobilized funding through the International Union for Conservation of Nature (IUCN) for a study on the challenges and opportunities presented by the possible use of "Big Data" and Artificial Intelligence (AI) systems for the management and conservation of high seas ecosystems.

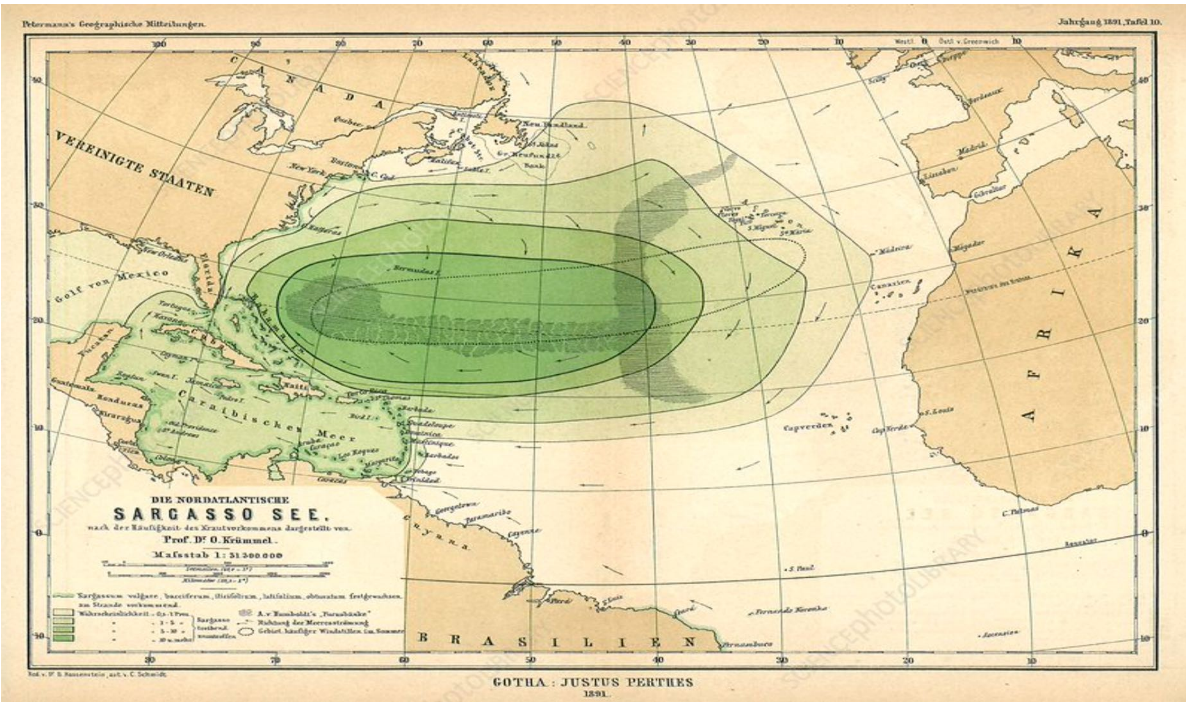


Figure 1: 1891 Sargasso See, Krummel Petermanns Lores (Wikipedia Commons)

2 Report Purpose & Methodology

The report[2] aims to articulate the potential opportunities in the utilisation of "Big Data" and AI in providing future ocean governance at the global level, set in the context of a high-level user requirements assessment of the challenges to effective ocean/high seas governance. Our research and analysis focused particularly on the complexities associated with ungoverned sea spaces in Areas Beyond National Jurisdiction (ABNJ), a key characteristic and fundamental challenge faced in the Sargasso Sea. It reviews the current state-of-technology in earth- and space-based remote sensing and the use of AI technologies to access and analyse Big Data, creating information in a strategic and cost-effective way for the governance, management, and conservation of remote ocean areas.

---

[1] Azores, Bahamas, Bermuda, British Virgin Islands, Canada, Cayman Islands, Dominican Republic, Monaco, UK & US.

[2] Caveat: this report has been somewhat restricted, both in timescale and resource available. As such, it does not seek to present itself in the style, or with the heft, of an authoritative peer-reviewed academic paper.

What the report does *not* aim to provide are specific solutions to the considerable barriers that exist in data-sharing; this is outside the scope of this work, although it does set-out the key data-sharing challenges that do need to be solved and makes some recommendations as to how some of these challenges may begin to be addressed.

To do this we drew on the advice of the SSC Expert & User Groups and others, to help us to assess these technologies and map potential future developments. We also leveraged the considerable capabilities of NLA International's in-house bespoke market intelligence curation and activation system – Blue Economy Knowledge System (BEKS) – to rapidly search on-line communities to assist in quickly building a contemporary picture of what is being discussed, researched, and operationalised. The report identifies key technologies, datasets and stakeholders including data and technology providers, associated existing and potential end-users, and assesses the possibilities and limitations of existing Big Data and AI capabilities and initiatives. It outlines some of the risks, challenges, and opportunities they present for effective surveillance, monitoring, and potentially enforcement, of conservation and management measures in remote areas of the oceans beyond national jurisdiction.

It aims to suggest ways in which small organisations such as the SSC might use Big Data and AI solutions to strategically influence the long-term data gathering, monitoring, valuing, governance and management of remote ocean spaces and their ecosystems, and suggest a user decision-making protocol, covering factors such as time, cost, quality, availability, and compatibility.

3 Executive Summary

Ocean Governance is a strategy used to manage human activities in an ocean, towards sustainable use and ecological regeneration. It is informed by, and includes, a whole range of economic, scientific, ecological, and financial activities and policies, covering all events in the ocean space, at local, regional, national, and global levels. The commercial use of the oceans in general (and especially of the high seas) can be seen as a tragedy of the commons. Almost two thirds of the oceans are 'high seas', often referred to as Areas Beyond National Jurisdiction (ABNJ) and are subject to limited Ocean Governance.

There is a growing appetite for increased governance of the high seas. Problems that start outside of EEZs can gradually migrate inside, both in terms of human activity, and its ecological consequences; modelling shows that damage to key high-seas ecosystems has as great an impact to coastal ecosystems as to the original locations. There is also a growing sense of inequity between those fishing inside and outside of EEZs; the accelerating and highly unequal industrial use of the oceans for mineral extraction, and harvesting of genetic resource, exacerbates these concerns.

A primary barrier to establishing ocean management in the high seas is developing sufficient, high-certainty evidence, to justify the implementation of policy. The requirement to evidence a likelihood of ecosystem damage from human activities is often interpreted as one to provide evidence damage already done, with a clear attribution of cause. This is an obstacle to taking a precautionary approach to industrial ocean use, and hinders early ecosystem protection, with a view to preventing damage until cause is better understood, and mitigations developed. Furthermore, establishing evidence with a sufficient level of confidence is an expensive and expert process – an ocean ecosystem assessment for a small Central American nation has costs of the order of six million US dollars. If the ambition is to assess and manage ocean ecosystems globally, we must seek solutions that reduce the cost of this process and improve access to the necessary knowledge and technologies. In part, this cost reflects a global inequity in expertise.

It also reflects the current state of 'Big Data' for ocean ecosystems; whilst vast quantities of ocean ecosystem data have been gathered over decades through both research and commercial activities, little of it has been practically operationalised with a view to ocean management. The state-of-play is characterised by inconsistent cataloguing, limited data sharing (for reasons of cost, difficulty, and perceived commercial/national sensitivity), poor standardisation (both within and across ocean sectors), and low data retrievability – requiring specialist expertise and comprehensive knowledge to find data and then select the right data for a particular task.

There is a pressing need to address these issues, and we propose four main avenues to do so: Promoting open-source non-rivalry data sharing for the high seas, premised on a global commons, justifying data transparency, and potentially seeking implementation in policy (e.g., through the Biodiversity Beyond National Jurisdiction (BBNJ) Treaty); Facilitating data sharing by all parties, including those already willing to do so (e.g., academia and NGOs), by establishing funding for data sharing (potentially in research grants and project budgets) and building expertise; Establishing data and meta-data standards across types and sources, towards maximising data use, discovery, and interoperability and, most importantly; Accelerating the design, realisation, and support of ocean Big Data sharing platforms, potentially aligned to the system requirements we put forward in Section 5.2.1, with a fundamental focus on end-user needs and non-expert use.

There is no likelihood that the scale of Big Data needs for ocean ecosystem analysis or management can be reduced; the complexity of the problem necessitates a Big Data approach. There is a requirement for high resolution 4D data, spanning large ocean zones to depth, over long-time scales, and from numerous remote sensing sources including earth observation, ocean-surface sensors, and those within the water column: the quantity and variety of necessary data is vast and increasing. There is also a need for targeted improvements to our sensing capabilities, substantial benefit could be derived from more sensors integrated on vessels/platforms of opportunity to provide ecosystem data and assist in observing human activity. In addition, there is demand for improved sub-surface sensing (particularly with a view to future industrial ocean uses, such as deep-sea mineral extraction). The necessary sensor technologies are all within the art of the possible, and the challenge is implementation, not development.

Establishing appropriate Big Data infrastructure is a pre-requisite to applying the many AI-based analytic tools that may be of benefit to the domain. Computing power, quantity of available data, and data gathering have historically been barriers to Artificial Intelligence based approaches. It is only relatively recently, with modern computational capabilities and data infrastructures, that these approaches have become viable. They present a natural evolutionary step in data analysis, modelling, and prediction – and as is already the case in many sectors, it is likely that they will come to underpin the methods used for ocean ecosystems analysis, and therein evidencing and implementing Ocean Governance. The numerous forms and benefits of AI are detailed in Section 5.2.2, but by way of simplification AI approaches fall into two categories. One is the intelligent automation of analysis, monitoring, and decision making, that is currently undertaken by trained human experts; this usually combines rules-based approaches (which codify human knowledge in algorithms) with 'machine vision' techniques such as image recognition (e.g., to identify vessels from satellite imagery). Therein, an AI system becomes capable of replicating expert practice, which is necessary to achieve the scale of analysis that comes with Big Data problems. Furthermore, if such systems are generally accessible, or to some degree built into data sharing platforms, this may be seen as a means of reducing the global inequity of expertise.

This category of AI use is predicated on *a priori* knowledge, it assumes we know what data is important, and that we already understand ecosystem behaviour, and can therefore codify its analysis. For specific, well-defined, tasks this may be the case (e.g., identifying vessels entering MPAs). However, for the general task of ocean ecosystem analysis it is not. Ocean ecosystems exhibit complexity; their properties are emergent, and exhibit deep correlations across physical, biological, climatic, and human processes. From the human perspective, piecing together all the information and correlations necessary to interpret and model this complexity is a near impossible task. Here, a second form of AI approach – using novel Deep Learning and Generative AI methods – could prove useful, which uses bias-free learning approaches to study all sources of data, free from human guidance or intervention, to identify these deep correlations, and then build models for system behaviour. These advanced approaches can assemble knowledge from data, allowing us to associate cause with phenomenon, explore hypothetical scenarios, retrospectively analyse past events, and better understand which data sources are most important. These may powerfully contribute to our understanding of ocean ecosystems, but, even more so than traditional AI, are dependent on a mature, highly consistent, Big Data infrastructure.

There is no doubt that Big Data and AI approaches have a key role to play in Ocean Governance, starting with scalable, cost-effective, evidence generation to inform management and policy. However, special attention must be paid to a crucial non-technical barrier to use – trust. The quality, or benefit, of a new solution is irrelevant if lack of trust prevents uptake. In purely operational domains trust can successfully be generated through demonstration (for example, the use of AI in various forms of autonomy); governance is not purely operational though – it has human, economic, and political consequences. Already, a barrier to establishing governance is building sufficient evidence; whether simply via data analysis, or through a complex, black-box, AI-driven system, the methodology for evidence generation must be trusted. Therefore, understanding and solving issues of trust are a priority. A perceived lack of trust can easily be used as a justification for inaction. Building consensus on what trust means, and how it is achieved, should run before and alongside any technical development – not after it.

The Big Data and AI concepts discussed within this report all represent the current art of the possible; they reflect cutting-edge practice in technologized sectors, which is now being tested in and translated to other domains. None of these solutions will be without cost. Big Data infrastructure is both costly to develop and to maintain, and AI solutions – whilst predicated on general concepts – will need to be tailored to, and trained for, the domain at hand. There is no surfeit of expertise in artificial intelligence, and many sectors are fiercely competing for limited capacity. As a priority, the technology needs for high seas governance should be formalised and communicated to the AI sector – as an issue of global significance seeking an immediate solution – ideally associated with clear, funded, pathways to feasibility testing and subsequent development.

3.1 Conclusions

- The provision of contemporary good Ocean Governance and the use of technologies underpinned by Big Data and Artificial Intelligence are inextricably linked.
- The Technology exists today to generate suitably diverse, relevant, and sufficient ocean data. This Big Data can be analysed using Artificial Intelligence to:
    - Generate the necessary understanding of the relationships between human activities and their impact on the complex ocean biological and environmental ecosystems.
    - Provide compelling evidence to establish the need for good Ocean Governance by informing decision-makers responsible for creating good Ocean Governance policies.
    - Generate convincing, near-real time, maritime domain situational awareness to enable policing and enforcement of human-related activities and where appropriate, underpin subsequent judicial action.
    - Provide suitable Measures of Effectiveness of in-place Ocean Governance policies to allow for subsequent review, revision, and release.

- A consistent barrier to justifying, evidencing, and implementing Ocean Governance is data sharing, availability, quality, and utility. Addressing these is a first order need and may be a pre-requisite to many AI methods. Specifically, data sharing and analytics platforms, combined with a drive for open data, is foundational to technologized governance.
- A key challenge in using Big Data and AI for Ocean Governance is one of trust; both trust in data and methods, and trust in key organisations and platforms to enable data sharing towards improved understanding. There are technical and human elements to this, and trust must be established alongside technology, with a focus on open data, algorithms, and methodologies.

3.2 Recommendations

- Facilitate Big Data standardisation, cataloguing, and the development of sharing and analytics platforms, paying equal attention to both technical and human requirements. Priority should be placed on technological ubiquity and equity, to improve cost-effective ocean ecosystem analysis, to support high-seas governance.
- Promote, incentivise, and support data sharing and FAIR data principles, and help to mitigate the costs of data-sharing that limit what is feasible, especially for researchers and NGOs.
- Initiate and advance discussions towards agreement or policy for open-source data from the high-seas, potentially using the Biodiversity Beyond National Jurisdiction (BBNJ) Treaty as a mechanism, with a view to opening-up data from private and national sources.
- Utilise Big Data and AI methods to distinguish the most important forms of data for ocean ecosystem analysis, and therein reduce extraneous data gathering, and enhance remote sensing for critical information.
- Explore and define specific use-cases for Big Data and AI-enabled/enhanced governance, including analysing service demand, cost, and acceptability, with a view to road-mapping the development of data and analytics services, identifying closest-to-realisation solutions, and pre-emptively addressing issues of trust.
- Strategically specify and then run feasibility studies for the use of AI methods for ocean ecosystem analysis, and to both automate and facilitate ocean management, seeking to demonstrate effective solutions and translate the state-of-the-art, from other technologized sectors.
- Develop technologically underpinned concepts for dynamic ocean ecosystem governance, using remote sensing and AI analysis to flexibly define protection as and when it is necessary, and to communicate this to all ocean users in a ubiquitous and accessible form.
- Undertake wider stakeholder engagement, seeking to understand what acceptable and good high-seas governance looks like, to both the industrial users of the high seas, and the many coastal economies whose livelihoods indirectly depend on high seas ecosystems, who need ecosystem analysis data products, and who may facilitate implementation. Engage in capacity building to improve the sectoral understanding of the state-of-the-art, and to improve access to Big Data or AI-based ocean management solutions.

4 The Needs of Ocean Governance and Technology

4.1 The need for governance

Ocean Governance is a strategy used to manage human activities in an ocean towards sustainable use and ecological regeneration. It is informed by, and includes, a whole range of economic, scientific, ecological, and financial activities and policies, covering all events in the ocean space, at local, regional, national, and global levels. The process of establishing governance should be granular, transparent, consultative, and ultimately evidence based. Ocean Governance necessarily involves action, response, and enforcement, requiring physical implementation at the lowest level, typically for remote sensing and responsive enforcement.

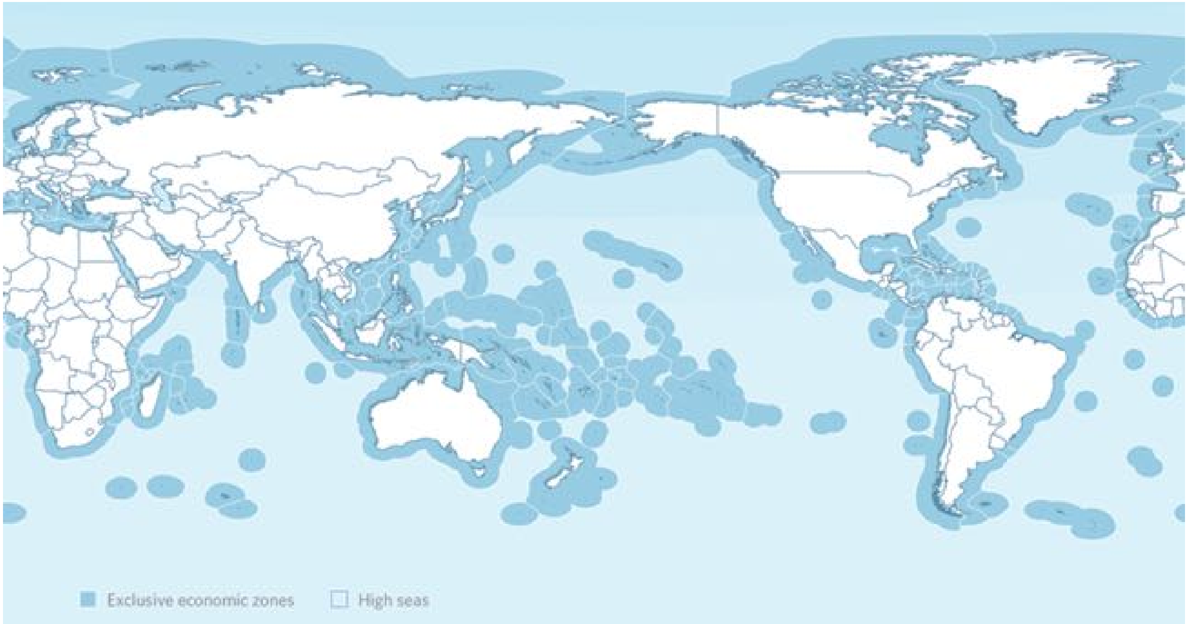


Figure 2: The light blue waters in this map represent all the high seas. ([Pew](Pew))

The commercial use of the oceans in general (and especially of the high seas) can be seen as a tragedy of the commons. Almost two thirds of the oceans are 'high seas', often referred to as Areas Beyond National Jurisdiction (ABNJ) and are subject to limited Ocean Governance. The United Nations Convention on the Law of the Sea (UNCLOS) regulates, to some extent, human activities in ABNJ, putting a premium on the "protection and preservation" of the ocean (Art 192), but the high seas regime has been called an "unfinished Agenda"[3]. When the treaty was negotiated in the 1970s many of the riches of ocean biodiversity and its resources were unknown. Moreover, the implementation of UNCLOS is varied – there are examples of good implementation (discussed later), but the prevailing view is that most implementation is poor and constitutes a race to the bottom exacerbated by entrenched interests and lacking transparency. Many discussions of Ocean Governance centre around commercial fishing and whilst it is not the only industrial human activity on the oceans, it is the one with evident ecological consequences, covering over 55% of the world's oceans. Fishing Fleets with increasingly global reach are being subsidised to fish greater and greater distances from their home shores. The consequence of this is diminishing resource availability and significant ecological damage, particularly to high-seas areas where the UNCLOS implementation has not

[3] Freestone, David. "Governance of Areas Beyond National Jurisdiction: An Unfinished Agenda of the 1982 Convention." UNCLOS at 30 (2015). Barnes, Richard, and Jill Barrett. Law of the Sea-UNCLOS as a Living Treaty. BIICL, 2016., pp. 231-266.

matched the threats posed by this activity or has only done so within a narrow remit (for example, there is rapidly growing, unregulated, squid fishing in the high seas, predominantly by long-distance fleets). In the absence of governance and strong implementation there is no reason to expect that behaviour will change without radical changes in the governance regime. Furthermore, other industrial uses of the ocean, such as seabed mining, whether in the high seas or EEZs (particularly of developing and industrialising nations), may well proceed in the same way if not subject to considered regulation.

Seabed. Photo: Kevin Clyde Berbano (Pexels)

In both economic and ecological senses, this practice is unsustainable. It can also be seen as self-reinforcing. In a tragedy of the commons, it is often the case that mitigating evident issues (such as resource depletion), and therefore enabling sustainability, is consistently seen as ‘someone else’s problem’. Whilst sustainable solutions and practices enabling long-term economic benefit might exist, they will usually require external impetus (often in the form of governance and investment) to change behaviour. This is not a theoretical issue, and we have seen collapses in fisheries already, such as the well-known case with Northern Cod. In general, we need approaches to managing our natural ocean resources more effectively.

One may comment here that a further challenge is the sector specific nature of governance and implementation; implemented by bodies that each have very focused remits. Ocean ecosystems are complex; ‘complexity’ is a term often used, but with a specific technical meaning – a complex system is inherently ‘more than the sum of its parts’, it cannot be characterised, predicted, or *managed*, by looking at each of its parts in isolation[4]. Therein, effective governance of ocean ecosystems may require cross-sectoral strategies, and governance frameworks that are designed with this in mind.

[4] This has had profound impacts to engineering and systems analysis in adjacent, complex, domains. For example, engineering complex semi-conductors for reliability has resulted in a ‘good + good = bad’ philosophy; by maximising the reliability of each component, the reliability of the system is dramatically reduced – it must be optimised “as a whole”. Whilst this may sound like an esoteric example, the nature of complexity is not domain specific; ecosystems must be treated as a holistic, complex, whole – not as a sum of independent parts. It may be that only by maintaining a holistic overview can ecosystem sustainability be achieved. One notes too that the recent (2021) Nobel Prize in Physics was awarded for the study of complexity, pertaining specifically to modelling global climate and therein climate change.

In areas with active Ocean Governance, efficacy can be reduced due to limitations in monitoring and enforcement. This is particularly true for Small Island Developing States (SIDS), for whom enforcement can be a challenge beyond immediate coastal waters. Remote sensing and analytics tools are seen as a potential solution here and have seen success in reducing illegal fishing around the Ascension Islands[5].

It is also not always the case that where governance exists, it is sufficient. Loopholes in governance, especially relating to fishing, drive exploitative practices. An example of this can be seen in high-seas squid fishing, which does not fall under the competence of Regional Fisheries Management Organisations (RFMOs) and is seeing rapidly increasing fishing, as recently reported by GFW[6].

There is a natural competition between governance and entrenched interests. For example, a challenge with limiting fishing zones is managing the displacement of fishing activities. Once the fishing fleets begin to regularly use a particular ocean space, moving them is challenging. Incumbent interests generally expect governing bodies to facilitate their continued activities post-displacement, which is challenging and slows policy implementation. Therein, there is a clear time imperative to implement Ocean Governance, particularly in areas where it is effectively absent. This has pertinence to the high-seas and in particular the Sargasso Sea.

4.2 Improving governance of the high seas and the Sargasso Sea

The points above hold true for Ocean Governance in general. However, there is an increasing focus on the high seas.

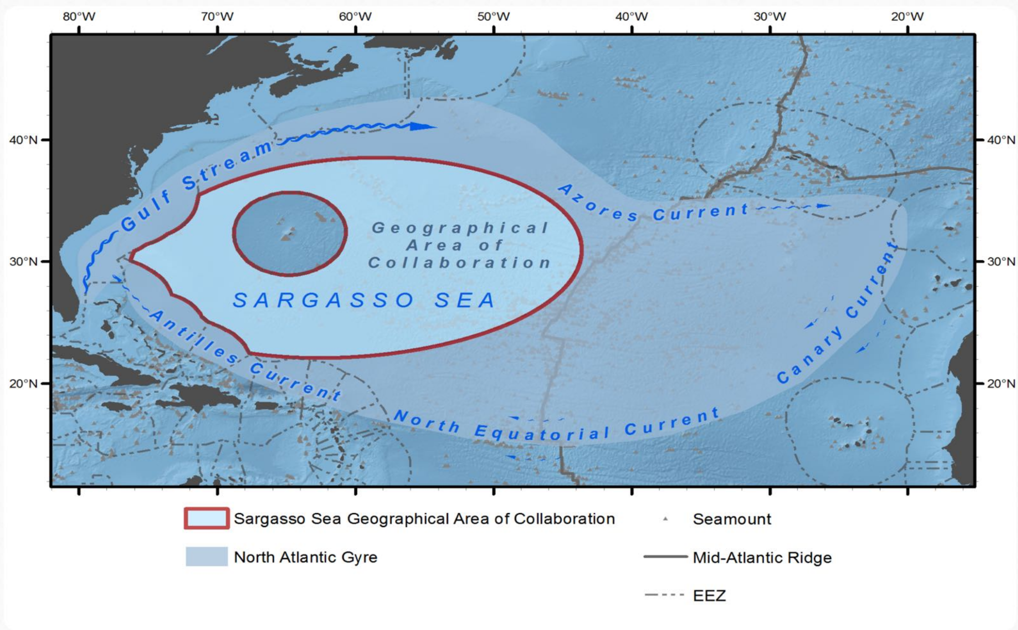
Tripletail in Sargassum. Photo: Lindsay Martin

There is a growing appetite for governance outside of EEZs. Problems that start outside EEZs can gradually migrate inside, both in terms of human activity, and its ecological consequences[7]. There is also a growing sense of inequity between those fishing inside and outside of EEZs. Very few nations have the vessels necessary for distant fish capture, activities which favour a minority of mainly developed nations, at the cost of global resource availability, and local opportunity[8].

---

[5] Rowlands, Gwilym, Judith Brown, Bradley Soule, Pablo Trueba Boluda, and Alex D. Rogers. "Satellite surveillance of fishing vessel activity in the Ascension Island exclusive economic zone and marine protected area." Marine Policy 101 (2019): 39-50.
[6] GFW: Squid Fishing SE Pacific 2020-2021 Seasons
[7] And current modelling shows that damage to key high-seas ecosystems has as great an impact on coastal ecosystems as at the location
[8] Sumaila, U. R., V. W. Y. Lam, D. D. Miller, L. Teh, R. A. Watson, D. Zeller, W. W. L. Cheung et al. "Winners and losers in a world where the high seas are closed to fishing. Sci Rep 5: 8481." (2015).

Thus, a key issue today is understanding how to collectively govern the high seas. Since 2004 the UN General Assembly has been discussing the issue of the conservation and sustainable use of biodiversity in Areas Beyond National Jurisdiction (ABNJ). In 2018 an Intergovernmental Conference (IGC) was established to negotiate a new international legally binding instrument on this issue. Whilst the negotiations have been interrupted by the COVID-19 pandemic, the hope is that the new agreement can be finalised in 2022.



Credit: Marine Geospatial Ecology Lab, Duke University.

Figure 3: Sargasso Sea Geographical Area of Collaboration

There is a perception by many involved in those negotiations that while there are existing international organisations which already have the competence to deal with the conservation of the high seas, they have not yet taken on a proactive role in responding to modern challenges. These organisations include the International Maritime Organization (IMO), as well as the Regional Fisheries Management Organisations (RFMO) established pursuant to the United Nations Convention on the Law of the Sea (UNCLOS), including the International Seabed Authority (ISA) and the Straddling Fish Stocks Agreement in 1995. IMO competences include all vessels and oceans, but it has been slow to react to the increasingly obvious problems of vessel impact on high seas ecosystems. There is a general need at the international level to catalyse action, and to enable governance and policy decisions.

One facet of enabling governance is the building of an evidence base of ecosystem issues, and credible solutions. This is what is currently being established for the Sargasso Sea, a two million square mile open ocean high seas ecosystem, and as such a very good case study for high seas governance. Central to this will be gathering and collating the data necessary to inform and justify any specific action or policy. The overarching method for this is to undertake an ecosystem analysis, which could underpin a scientific approach to ocean management and governance, and lead to actions such as defining Area Based Management Tools (ABMT), including possible Marine Protected Areas (MPA) underpinned by coherent Marine Spatial Planning (MSP). It should be noted that whilst most modern legal instruments mandate the use of a 'precautionary' model of governance, establishing policy early based on the risk of negative effect, in reality the burden of evidence appears high.

There are a range of outstanding issues that the Sargasso Sea Commission have identified that require prioritised governance responses. Fishing for tuna and tuna-like species is within the competence of the International Commission for the Conservation of Atlantic Tunas (ICCAT), and demersal stocks above 35°N are regulated by the Northwest Atlantic Fisheries Organisation (NAFO). However, fishing for other species is currently not regulated at all. Data from Global Fishing Watch (GFW) shows that fishing activity is increasing in the Sargasso Sea, but it is neither evident what is being fished, nor the impact of these activities on the ecosystem. The same is true for the prospect of future deep seabed mining in ABNJ. The ISA have granted exploration licenses for three prospective sites on the Mid-Atlantic ridge adjacent to the Sargasso Sea, but the consequences of potential activity on the surrounding ecosystem are not understood. Plumes from future mining activities could easily sweep through the area, causing unpredictable but significant damage[9].

Atlantic Bluefin Tuna. Photo: Richard Herrmann

4.3 The relationship between policy, governance, and technology

The relationship between governance and technology is not simple. There is an interplay whereby innovative technology helps to formulate and justify policy, as much as to implement governance. Ocean ecosystems, including human activities, are evolving and complex. Our growing ability to interpret, predict, and monitor these may enable highly nuanced governance that is aware of, and responsive to, the state of the ecosystem. For example, adjusting dynamically, based on the migration of fish stock, accommodating changing locations for spawning grounds, and optimising ecosystem regeneration. Ultimately this could lead to more targeted restrictions, based on a deeper scientific understanding, and substantially improved sensing and monitoring.

Therein, whilst governance decisions (captured in a legal and regulatory framework) are a pre-requisite for the use of innovative solutions, demonstration of the art of the possible may be necessary to catalyse decision making. A challenge here is the investment case; advanced solutions, such as those built on AI, may be costly to develop, and only economically viable in the context of long-term services. However, their development may be necessary to explore

[9] Recent modelling of the Clarion Clipperton Fracture Zone in the equatorial Pacific indicates that plume dispersal from a single mining operation could cover 1000km over a year, see: Muñoz-Royo, Carlos, Thomas Peacock, Matthew H. Alford, Jerome A. Smith, Arnaud Le Boyer, Chinmay S. Kulkarni, Pierre FJ Lermusiaux et al. "Extent of impact of deep-sea nodule mining midwater plumes is influenced by sediment loading, turbulence and thresholds." Communications Earth & Environment 2, no. 1 (2021): 1-16.

the forms technologized governance can take, and to establish the evidence necessary to generate action. It is clear from past cases that the introduction of governance can spur wider investment (e.g., water treatment technologies to reverse eutrophication in the Black Sea, and the new innovative industry in ballast water treatment solutions born of GloBallast) but, until that point, there is a need to enable and de-risk technology realisation.

Central to these matters is the need for data, both as a form of evidence, and an underpinning of AI/analytics technologies. Whatever the use, there is a need to maximise free and open access to data in and of itself. Data access and sharing is a policy-level issue involving public, private, and scientific stakeholders; currently data availability and utility varies significantly. Addressing this may be a pre-requisite to effective high-seas governance. The details of data utility will be discussed later, but there is a general lack of data standardisation within and across ocean industry and science domains, perhaps apart from satellite data. There is also a lack of consistent meta-data, tagging and organisation; even if relevant data already exists, discovery and retrieval can be extremely challenging.

Photo: Cottonbro (Pexels)

More fundamentally, there is a need to encourage data sharing and open science. Data transparency is a primary issue; many RFMOs, states, and private bodies are reluctant to share data, the justification for this is commercial confidentiality, and risk to competitive advantage. However, the view by many in the domain is that this data obfuscation prevents real scrutiny, and – for the global ocean commons – this is deeply unacceptable. Mechanisms to enshrine data transparency and sharing in high-seas agreements could be a major enabler for ecosystem analysis and sustainability. In the scientific domain there is no view that data is strategically withheld, rather there is a time and fiscal cost to data sharing, and a technical challenge. Agreeing platforms for data sharing, and incentivising organisations to make their data available, is a priority. In the context of science, a minimum standard of data sharing – perhaps following sector agreed guidelines – could be built into future grants to ensure the effort is made to share collected data. The cost of doing this should not be ignored, scientific organisations operate at the limits of their budgets, and a model of financing long-term data availability may be needed.

In the context of analysing and collecting data on the human and industrial uses of the oceans, there is no sense that the barriers are predominantly technical. Current capabilities for tracking and monitoring vessels are generally sufficient already, and AI analytics are being successfully used to determine if vessel operations are appropriate or not. There is a need to improve verification, validation, and ground truthing, especially in the context of unintuitive AI methods, and building trust in derived data products. Emerging technologies can improve this, and help with scalability, but again the core barrier is sharing of high-quality vessel tracking and catch data; this includes the meta-data necessary to ascertain quality, such as attribution, traceability, and quantified uncertainty – all of which are discussed further in Section 5. Therein, the crux might be to encourage a culture of transparent operations, with fairness of ocean use guaranteed through mutual visibility, at least insomuch as each stakeholder has confidence others are not misusing the ecosystem. Any steps towards this are to the benefit of governance, and transparency weakens the influence of entrenched interests on policy decisions.

The challenge is not just data sharing; it is putting the right data and information in the right hands, enabling decision making and policy formulation. Understanding the form this needs to take, the enabling tools, and how to strengthen the translation of scientific advice to policy, is a priority challenge. Analytics and AI may play a key role in translating multivariate, layered, data into *synthesised holistic metrics* that are interpretable across decision-making and industry. Raw data, however well-structured or layered, lacks the interpretability necessary for wide use, particularly considering the inequity of knowledge in ocean science and data analysis. Data are fragments from which information and knowledge may be derived, it is the latter that should be presented by an accessible data product. The process of moving towards technologized, data-driven, governance needs to be road-mapped; exploiting the two-way relationship between technology and policy is an opportunity for transparency, objectivity, and better decision making.

4.4 The role of technology in Ocean Governance

Having considered the need for and the nature of Ocean Governance in the high seas, it is necessary to explore the specific role of technology. What it is that we are looking to achieve through Big Data, AI, and remote sensing solutions, and therein the technical requirements that prospective solutions must meet.

The issues discussed in the previous theme on policy and technology are echoed here. There was a view from experts and end-users that we will need to work backwards from management and governance frameworks to specify what technical approaches are necessary, and then to understand constituent data requirements and innovative solutions for each case, necessitating the development of policy as a starting point. There was no consensus on this point, and irrespective of it, several technology ‘needs’ were apparent. Broadly the role of technology in ocean governance falls into four categories:

- Establishing the evidence necessary to inform and justify policy decisions.
- Enabling and implementing scientifically based management.
- Implementing enforcement, and
- Deepening understanding of ocean ecosystems – which is indirectly related to governance by improving the scientific foundation and developing models.

Some context is also provided by the Ocean Innovation Challenge, which is an accelerator for emerging solutions. It is funding innovation pertinent to the UN's Fourteenth Sustainable Development Goal (SDG14), to conserve and sustainably use the oceans, seas, and marine resources for sustainable development. So far it has funded technology towards marine pollution reduction, and sustainable fisheries.

Considering the first category; establishing evidence relies on gathering and collating data (across ecological, economic, scientific, and industrial domains), deriving from it a snapshot of the ecosystem and its activities, identifying key risks and issues (especially ecosystem threats, such as loss of biodiversity due to vessel traffic), and predicting its evolution with and without proposed governance. This relies on a suite of technologies including: the full gamut of earth-observation, surface, and sub-surface remote sensing capabilities for data gathering; the Big Data and data sharing implementations that allow for this multi-modal information to be stored, retrieved, and utilised; and data processing, analytics, and insights tools that make sense of the information. These analytic tools could include AI methods (of various sorts, from mature rules-based approaches to more nascent techniques such as reinforcement learning and generative AI), but these will likely be adjunct to existing, successful, statistical methods and models.

Eel Leptocephali. Photo: Marko Freese

The UNDP has a formalised methodology for implementing governance, which starts with a trans-boundary analysis covering much of the above. This process usually relies on historical data; whilst there have been some exceptions, such as a twelve-million-dollar oceanographic assessment in the Indian Ocean, typically the budget for formulating and justifying policy is not sufficient to enable new data gathering. Establishing governance directly depends on data first and foremost.

The general view is that a lot of data exist, covering many different and subtle aspects of ecosystems; the sum knowledge from decades of marine research is substantial. However, much of these data are not operational. There are several barriers to operationalisation: much of the data lack standardisation (in multiple ways, including the temporal domain, depth regimes, ocean gridding, and data format) introducing challenges for interoperability; quality of data, in terms of resolution, coverage, and timespan, varies greatly; data certainty is not consistently expressed; and – perhaps most importantly – the right data sets can be difficult to find. For common ocean properties there can be unmanageably numerous data sets, each encumbered by choices specific to their data gathering process, and not necessarily suitable for all purposes. It is practically challenging for ocean managers, and non-data scientists, to understand what a good choice of data looks like for their purpose. This is doubly significant to developing nations, for whom there may be a substantial knowledge and expertise gap.

Addressing these problems is an acknowledged priority, with projects such as NASA's COVERAGE[10] currently looking to do so. Given the enormous variety of data sources and types associated with ocean ecosystems, and the great variety of stakeholders holding data, it is extremely unlikely that a single platform approach would be suited to this domain. There is a need to understand "what good looks like" for Big Data sharing and analysis services, seeking to maximise accessibility and broad usability, whilst providing an easy, low-cost way for data gatherers to open their data to public use (we articulate an idea of 'good' in Section 5.2.1).

Humpback & Sargassum. Photo: Andrew Stevenson

When it comes to implementing governance, one key outcome is Marine Spatial Planning (MSP). MSP is about documenting existing sea uses and ecosystem properties, developing a rich picture of all activities in the region and their interactions. With the variety of physical, biological, and human activities in an ocean space, this problem is manifestly complex. A complex system can be characterised as being 'more than the sum of its parts'; practically this means that understanding each facet of an ecosystem in isolation is not sufficient to understand how it will evolve as a whole – with emergent properties developing due to the many unconstrained, interrelated, processes within it. Understanding and managing complexity is more than a Big Data problem; AI methods have proved their worth in interpreting and modelling complex systems in other domains (including climate physics, but also smart cities, transport networks, and more). Translating these approaches for ocean governance and MSP could improve ecosystem understanding, help to analyse what the most important data types and sources are, and – by translating layers of data into a holistic, predictive, overview – provide operational knowledge to end-users. Currently an MSP exercise for a small central American country, costs in the order of six-million-dollars; globally MSP would cost billions, and so the use of technologies to bring this cost barrier down is necessary. A deeper predictive understanding of ocean ecosystems may also enable more responsive and targeted decision making to the benefit of ecology and economics.

[10] The CEOS Ocean Variables Enabling Research and Applications for GEO (COVERAGE) initiative is a NASA-led research and development project and cross-cutting, collaborative effort within the Committee on Earth Observation Satellites (CEOS) that seeks to provide improved, more seamless access to inter-agency, multivariate satellite data spanning the four CEOS Ocean Virtual Constellations – sea surface temperature, ocean vector winds, ocean surface topography, and ocean color radiometry – in support of ocean science and marine resource management applications for societal benefit (https://coverage.ceos.org/overview/).

As much as the role and impact of Big Data and AI technologies may have on future Ocean Governance, the barrier to access must be considered carefully. Global problems require global solutions; candidate technologies must be scalable and have pathways towards ubiquitous use. In the context of data technologies neither of these barriers should be fundamental, but it is important that technology developers are as aware of these requirements, as of the technical challenge. From the perspective of catalysing better governance, developing new solutions is as important as reducing financial, capacity, and expertise barriers, and ideally all should be pursued in synergy.

So far, this analysis has centred around data sharing and data analysis in one form or another. There is also a need for more, and better, sensing. Our understanding of ocean physics is increasingly competent, and it is characterised by parameters that are comparatively easy to sense, with ocean-surface (and shallow sub-surface) characteristics being observable from space, and deep-water observations requiring in-situ sub-surface sensors. However, ocean biology is much harder to sense, model, and understand. As is the influence of the physical environment on the biological, such as the effects of climate change on fish stocks and migration. Ocean and climate physics are not entirely understood either. For example, understanding the fronts, gradients, and air-sea coupling that contribute to hurricane formation is not solved. Furthermore, human activities such as seabed mining occur in areas we know little about, with consequences (ecological and physical) that we do not fully understand. Therein, the data needed to understand ocean ecosystems is fundamentally '4D' in nature; requiring coverage of the oceans through the water column, over long timespans to understand eco-system change and behaviour, but also at sufficient resolution to identify specific activities (e.g., illegal fishing). Improving our understanding will require more sensing, better sensing, and improved analytics – meeting data needs would require a significant enhancement to current capabilities.

Sargassum. Photo: JP Rouja

Ultimately, a system-of-systems encompassing improving data gathering, communications, sharing, processing, and analysis, will be to the benefit of any of the aforementioned challenges, and most certainly improve and lower the barrier of entry, to Ocean Governance. However, the path there will be subject to fiscal constraints, and necessitate prioritisation. The most pressing need is ubiquitous data-sharing and standardisation, providing utility to the ocean managers and non-expert end-users, not just data scientists. Therein, Big Data solutions should be a sector priority.

4.5 Examples of good governance

In the interviews and questionnaire responses several examples of ‘good’ ocean governance were highlighted as case studies of what can be achieved, they are presented here to provide some context, and shed light on the relationship between policy decision and investment in solution development.

A pertinent example of a technology driven solution for illegal fishing can be seen in the ocean space around the Ascension Islands. Satellite surveillance and analytics from Ocean Mind were used to help monitor fishing activities. On the enforcement side, very high fines were implemented for those who fished illegally. Knowledge of this system was pro-actively promoted; once the fishing community understood that a capable monitoring system was in place their behaviour changed and illegal fishing was substantially reduced. Furthermore, the implementation of better monitoring enabled the detection of a different problem; vessels transporting dangerous cargo. Rowlands et al, 2018 details this case[11].

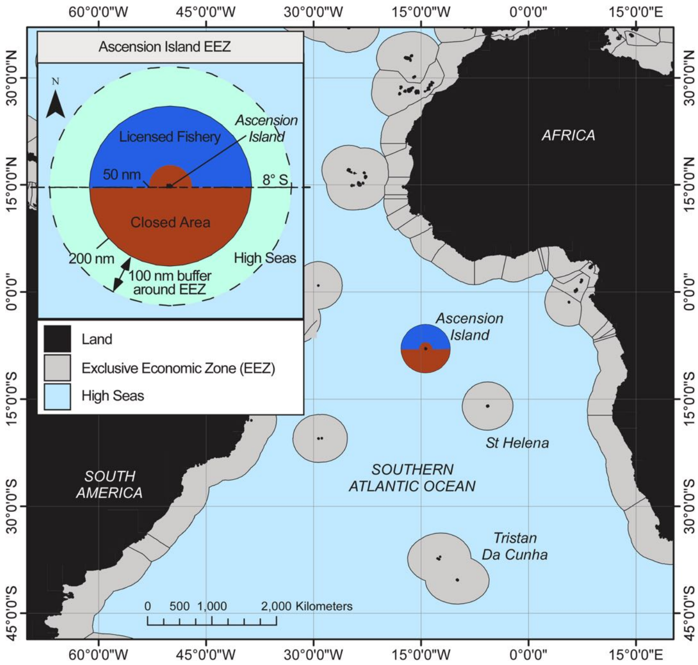


Figure 4: Ascension Island EEZ, Southern-Atlantic Ocean.

In the former case extant technology was applied to test its impact; more often policy is set, and technology then develops to meet new demands. An example of this can be seen in the ‘The GloBallast Story’. A primary cause of the migration of invasive species to new waters has been attributed to the ballast water of long-distance vessels. These waters carry species across the oceans, at times resulting in transference to new waters where they destructively thrive. To combat this a global convention on ballast water was negotiated, adopted in 2007, and put into force in 2017. It provides technically specific stipulations on how clean ballast water must be to avoid transference of species. This clear technical guidance combined with enforcement methods has led to substantial growth in ballast water cleaning technologies, which are now widely operational. The drive for competitive compliance solutions has generated a $40bn industry. Whilst this is neither data nor AI, it is an example of successful technology-oriented governance, in this case driving innovation.

[11] Rowlands, Gwilym, Judith Brown, Bradley Soule, Pablo Trueba Boluda, and Alex D. Rogers. "Satellite surveillance of fishing vessel activity in the Ascension Island exclusive economic zone and marine protected area." Marine Policy 101 (2019): 39-50.

Similarly, policy driving innovative solutions can be seen in efforts made to reverse the eutrophication of the Black Sea. Policy on water treatment, combined with innovation funds, led to a new burgeoning industry. This provided substantial economic benefit to the region, whilst addressing the key issue of eutrophication.

Technology aside, several agreements are held up as examples of successful Ocean Governance and may inform governance in the high seas. In interviews, the OSPAR convention was most often cited as an example of successful governance, providing marine protection and ocean sustainability through regulatory agreement between ocean bordering states. The Convention on the Conservation of Antarctic Marine Living Resources (CCAMLR) is similarly seen to be effective in regulating fishing and has established key MPAs.

The Western and Central Pacific Fishing Convention (WCPFC) provides a good example of a multi-country regulatory framework that served to operationalise the UN Straddling Stocks Agreement in the western and central Pacific. It has been highly successful, reducing overfishing of all four tuna stocks in this area (representing over half of global tuna yields) to 6%, resulting in complete sustainability for these fishing stocks. Practically, this regulation was enhanced and enforced via technologies: Vessel Monitoring Systems (VMS) and on-board observations systems were mandated for all vessels fishing in the waters. With these monitoring capabilities available, the vessel day scheme was implemented, auctioning off daily fishing rights – economically enhancing the participating states, a boon particularly appreciated by the regional SIDS. Here governance was made achievable through technology that dramatically increased capacity for compliance, monitoring, and enforcement.

Most examples of successful regulation have been achieved top-down. Partnerships in Environmental Management for the Seas of East Asia (PEMSEA) is a counter example, which was developed bottom-up, and involved agreement across east Asian states. Its signature result has been the introduction of Integrated Coastal Management (ICM) with cross-sectoral planning. By developing the methodologies and tools to do this, they have scaled up ICM in east Asia from close to none in the 1990s, to covering about 40% of the east Asian coast today. Governance and regulatory frameworks have been put in place to achieve this, generally seen at municipal and provincial levels. Implementation of this has involved Marine Spatial Planning, the cost of which has largely been taken on by local governments; undoubtedly there is scope for future Big Data sharing platforms and intelligent analytics to improve capability and drive cost down.

5 Technology for Ocean Governance

There are several key roles for technology in creating and sustaining good Ocean Governance. This involves the collection and analysis of multi-source data, to produce evidence to inform Ocean Governance policy makers of where and how to act. Data are also used to inform and direct enforcement activities and can subsequently be used to measure the effectiveness or otherwise of the governance measures put in place – this is a technology-enabled data-cycle.

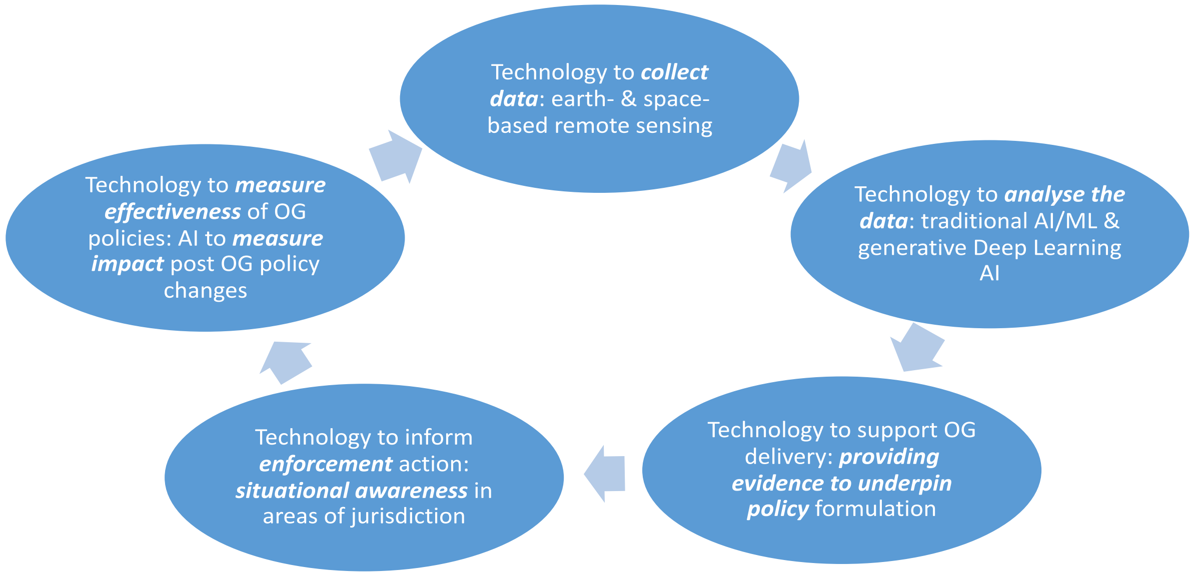


Figure 5: Technology-enabled data-cycle to create and sustain good Ocean Governance

5.1 Data Collection, Remote Sensing & Enforcement Technologies

5.1.1 The data journey so far

Data gathering in the world's seas and oceans has been ongoing for centuries. The first recorded scientific data collection in the Sargasso Sea is from the middle of the 19th Century and it has continued ever since. With the advent of steam ships and the industrial revolution, maritime routing and access to the Sargasso Sea has increased. Mariners are no longer limited by the vagaries of the complex maritime environment, particularly reliable wind, which was historically problematic in transiting the Sargasso Sea. The way in which data on environmental, biological, and human activity is collected has also changed considerably. Once only possible by the physical presence of an observer on a vessel, today technology-enabled remote sensing, both planet- (land and sea) and space-based, make it possible for a multitude of data relating to many differing characteristics and activities, environmental, biological, and human, to be collected 365-days-a-year, regardless of location, time of day or weather conditions.

The most recent large-scale scientific examination of the Sargasso Sea took place a decade ago which created an impressive amount of evidence relating to the existing ecosystem and the potential impact of human activity. The enhancements in data collection since then, specifically but not exclusively remote space-based sensors, alongside considerable advances in computing power and AI, (both rules-based Machine Learning, and the recently emerging Deep Learning, generative AI) mean the sheer scale of evidence available for policy makers to consider has increased manifold.

5.1.2 Technology to measure and address human-behaviour

Data collection relating to the environment and to biological activity does remain extremely challenging, but it does not have to deal with the purposeful obfuscation of certain types of activity, practically unique to the human species. Whilst fish and other sea creatures do seek to hide their whereabouts, they do this to evade likely predators, driven by a Darwinian instinct to avoid being eaten. This is contrary to the human predilection of doing so to gain unfair advantage and achieve some form of personal gain regardless of others, and with no thought for the morality and future sustainability of their actions. Without delving into an overly philosophical discussion at this point, it is worth highlighting the ever-increasing positive role technology can play in this seemingly unending battle against the human wish to operate in the shadows or darkness of the world's oceans, to amass ill-gotten gains.

5.1.3 The Electromagnetic Spectrum in remote sensing

The electromagnetic (EM) spectrum has played a vital and increasingly important part in quite literally, shining a light on human maritime activity. The advent of Radar in the Second World War and the considerable tactical, operational and safety benefits it provides, both in the air and maritime environments are well-documented. More recently, a different part of the EM spectrum has been used to further improve the safety of maritime shipping; radio-borne data from the Automatic Identification System (AIS) is now mandated to be fitted to ships over 300 tonnes displacement to allow suitable safe separation from each other. But this equipment (like VMS mandated for industrial fishing fleets by some nations), suffers from 2 key weaknesses when in the hands of operators engaged in illicit activities; it can be modified to show incorrect information (so-called *spoofing*), or it can simply be turned off (creating so-called "dark-vessels"). Both these practices can in themselves be key indicators to law enforcement organisations of *potential* illegal activity, but it is just an indicator, it is not evidence.

5.1.4 Space-based sensors

5.1.4.1 Essential Ocean Variables.

A key contribution of satellite remote sensing is on the routine, broad-scale monitoring of essential ocean variables such as sea surface temperature, surface salinity, ocean surface topography, ocean primary production, sea ice extent, water mass volume and related, derived variables enabling identification of dynamic oceanographic features, e.g., eddies, fronts, etc., and long-term trends e.g., global mean sea level rise. Much of these data are assimilated operationally into numerical weather and ocean models and are also the basis of climate modelling work.

Photos: Pixabay (Pexels)

5.1.4.2 Vessel Monitoring.

AIS was designed as a short-range information sharing platform, but it is now also mounted on numerous communications satellites extending its reach globally. This adds to other rapidly improving space-based remote sensing capabilities which are making a real difference today. Space was once the unique domain of the security and defence organisations of developed, powerful and wealthy nations, but now highly detailed, time-sensitive data from a multitude of space-based sensors can be collected and sold, or simply bought, by commercial entities. As a result, when a vessel's operator decides to turn-off their AIS or VMS transmitter to hide their whereabouts and activity, there are now an array of other sensors already on hand to continue to illuminate the situation – "sea blindness" can be turned into "sea vision". Satellite-based Electro-optical (EO) cameras, the Visible Infrared Imaging Radiometer Suite (VIIRS) and Synthetic Aperture Radar (SAR) arrays now produce very high-definition imagery of ever smaller sized vessels. With ever-reducing revisit times measured in hours not days, enabled by more complex but more sustainable non-polar orbits that keep the satellite in permanent solar view, thus providing constant power to the onboard batteries, whilst also allowing satellite taskers to concentrate their sensor time on areas of the oceans with the most human activity, these capability improvements make even the furthest oceans a difficult place to navigate completely unseen.

5.1.4.3 Specific Emitter Monitoring

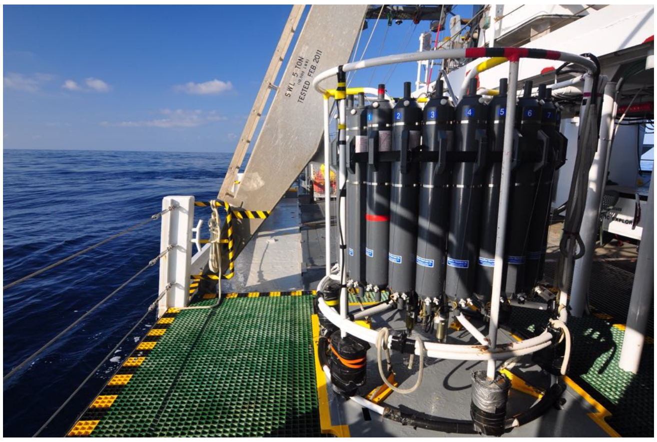
But even SAR, EO and VIIRS are not omniscient. Whilst SAR is not affected as much as EO and VIIRS by poor weather conditions and time of day, it still does not provide the incontrovertible evidence often necessary to move to a court of law. EO can provide this in the best weather conditions and satellite orientation, but recently a relatively old technology that is becoming more widely available is EM frequency transmitter fingerprinting (in military circles this is referred to as ELINT – electronic intelligence). Put simply, each transmitter on a vessel (e.g., radar for navigation, V/UHF radio for ship-to-ship communications, satellite telephone for speaking to vessel owners/accessing the internet) has a unique frequency fingerprint, meaning that whenever or wherever it is turned on, if it is "in view" of a suitably configured detector mounted on a satellite, its position can potentially be determined. This sort of system can also be terrestrially, or aircraft/drone based, although this will clearly reduce the range of detection to vessels operating in coastal waters or perhaps to EEZ boundaries; it will not cover the more distant high seas and ABNJs. On its own this fingerprint detail might not be sufficient, but when fused with multi-source data from other planet- and space-based sensors, a complete and compelling evidential picture can be created and presented to appropriate law-enforcement agencies for further action. As mentioned earlier, and although ELINT was not used in this specific case, when the Ascension Island fisherfolk who were engaged in illegal fishing became aware that certain monitoring technology and capabilities were now available in their region, they changed their illicit behaviour.

5.1.5 Underwater remote sensing and human activity

Data gathering of the oceans is not only carried out by space-based sensors. To fully understand the oceans, data must be gathered from within; put simply, sensors must "get their feet wet". Due to the physical characteristics of water, the capability to monitor using EM energy is limited to the space above, on, or just slightly below the sea surface. But this limitation can be overcome by exploiting the unique characteristics of sound travelling in water. Sound is particularly useful for tracking and understanding biological activity and other natural undersea phenomena, but it is also having an ever-increasing role in the tracking and recording of human activity – for example the use of hydrophones to listen for human-related insonification, either accidental or deliberate, which can have such a detrimental impact on marine life. But, where the near-ubiquitous nature of space-based platforms can be used to illuminate the underwater domain, is in the receipt and retransmission of the wealth of underwater data being collected by these sub-surface arrays. The use of communication satellites to relay data from remote oceans to remote laboratories, especially when near-real time transmission is needed, is fundamental, thus allowing this wealth of data to contribute to providing the evidence necessary to support good Ocean Governance.

5.1.6 In- and On-ocean remote sensing and data gathering of non-human activity

In-ocean remote sensing tools that measure physical factors are much more effective than those focused on measuring biological factors; to compound this challenge there are also orders of magnitude greater variability in biological sensing. Furthermore, biology and climate physics have additional but potentially very different sensing requirements to those for monitoring and managing human activities. Basically, understanding the biology and the underwater sensing picture, is more challenging than measuring and understanding human activities. More positively, from our research it appears that contemporary remote sensing capabilities are already highly capable, but they are constrained by the lack of commercial demand, and the practical challenges of monitoring vast ocean spaces. The problem of in-ocean remote sensing can be broken down into three distinct themes: understanding the physical and biological processes in the ocean – i.e., what is happening; monitoring natural activities, and monitoring human activities.



Bermuda Atlantic Time-series Study. Photo: Bermuda Institute of Ocean Sciences

Alongside the need for sufficient monitoring to understand what human activities are taking place, and how they are affecting the ocean, there is also a need to monitor the essential ocean variables. As mentioned above, satellite-based sensors can detect and measure surface information such as: temperature; roughness; salinity; acidification; and human activity, but this is not sufficient to understand the complex physical, biological, and human processes taking place within the entire water column and on and under the seabed. Sub-surface observation systems can look at essential ocean physical and bio-geochemical variables such as chlorophyll and turbidity, and biological variables such as, where life is, what lives where, as well as human activities. Sub-surface data is crucial, but not yet sufficiently available.

Passive ocean sensors are improving and moving towards reducing the ecological and environmental costs of their presence. Passive drifters, low energy gliders and sail drones, as well as high-altitude, very long endurance airborne drones are much less carbon-heavy assets that also do not require as much care and maintenance. Furthermore, so-called "platforms of opportunity", e.g., fishing vessels, cargo vessels, leisure craft and ferries moving through the oceans as part of their normal business, could be used to gather a wealth of data, possibly regulated in some way to "pay-back" their carbon usage. It is highly likely that many maritime users would be willing to do this, but once again, the requirement to decide data standardisation and types well in advance needs addressing. There are several ongoing programmes that are seeking to increase this type of "non-human" ocean observance such as The Global Ocean Observing System[12], the Global Ocean Acidification Observing Network[13] and Go-Ship[14].

NOAA and Saildrone Inc. are piloting 5 saildrones in the Atlantic Ocean to gather data around the clock to help understand the physical processes of hurricanes. Photo: Saildrone Inc.

5.1.7 Affordability vs Cost Effectiveness

It must be noted, that much of this data does remain expensive and beyond the budgets of many small or developing nations, but it is becoming more affordable. Arguably, when you consider the long-term detrimental impact of much of the maritime human activity and the costs of recovery and regeneration, a case could easily be made for buying data to build a case for good Ocean Governance, rather than dealing with the downstream impacts, is much more cost-effective. Affordability of data is perhaps a challenge that lends itself to regional, inter-governmental, international, or even philanthropic cooperative solutions to solve.

---

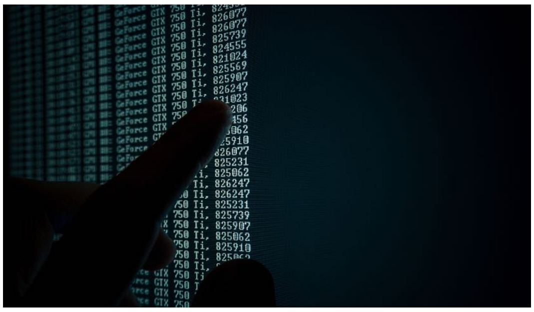
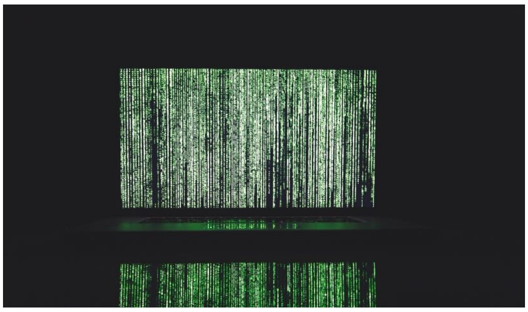
[12] Global Ocean Observing System
[13] Global Ocean Acidification Observing Network
[14] Go-Ship

5.2 Big Data, AI and Analytics

5.2.1 The Needs for Big Data Solutions

As we have discussed, a consistent barrier to justifying, evidencing, and implementing ocean governance is the availability of evidential data and analytics. Establishing a Big Data picture is part of building the evidence necessary to justify policy; as shall be discussed, this is currently challenging and costly. There are a variety of needs for, and ancillary to, using Big Data in this domain, broadly they centre around data sharing, availability, quality, interoperability, and utility.

These needs are as technical as they are human; it is crucially important that Big Data is not seen as the unique domain of the data analyst, computer scientist, or technical practitioner. Whilst these may architect and implement solutions, the end-user is necessarily the Ocean Governance community, with varied specialisations, a focus on ocean management and policy, and ranging access to tertiary knowledge and expertise (from nations at the forefront of technologization, to SIDS and other developing nations). There is a need for technical development in this domain, to better assist establishing governance in the Sargasso Sea or elsewhere, but a sufficient solution must be ubiquitous, intuitive, and accessible. It is incumbent on those developing solutions to minimise barriers to use. Therein, it is also important that the Ocean Governance community is an active stakeholder in all ongoing development, ensuring that their needs are well understood and well addressed – this can already be seen to be the case in the relationship between the Sargasso Sea Commission and NASA/JPL with respect to their COVERAGE project.

Photos: Markus Spiske (Pexels)

Addressing these needs towards enabling Big Data for Ocean Governance is a first order priority and is a pre-requisite to many data-hungry AI methods and analytics. It is also a pre-requisite to us understanding the state-of-play – earlier we posited that it is a challenge for ocean managers to select the right data for their needs from the vast variety and quantity available. More abstractly, the challenge is understanding what 'good' looks like for data, for a specific purpose. Sister to this is knowing how much of this is what we already have, and are collecting, and where we need to improve, or add to, data collection.

These questions can only be addressed in knowledge of the whole; the data needs to be catalogued and shared. Furthermore, it must be comparable, invoking a need for standardisation across disciplines, sub-sectors, and collection methods covering at least: meta-data; formats; metrics of certainty; temporal information; depth regimes; and ocean spatialisation/gridding. It is only when the whole can be seen, and parts measured against one another, that we can understand – through human processes and analytic algorithms – the strengths and shortcomings of what is available.

Standardisation is also a pathway towards data interoperability and compatibility. Systems utilising Big Data for evidence, governance, and monitoring, must not be brittle to technical change. As ecosystems and human activities evolve, along with our understanding, such systems will need more and different varieties of data, they may need better data in specific cases (or, perhaps, be able to make cost savings and do with less in others), data from new and changing sensors, and data from different providers as projects and businesses change. Unless its purpose is highly specific a rigid, brittle, data system is unlikely to be long lived. Standardisation, with a focus on data findability, interoperability, compatibility, shareability, traceability, and comparability, is a necessary enabler for the type of Big Data solutions that are needed. Many of these points have been formalised, in general, under the FAIR data principles[15], which could be applied directly to ocean data, and particularly to data collected on the high seas.

Photos: Mati Mango & Nátalie Rodrigues (Pexels)

The technical needs associated with Big Data also deserve mention, typically described as volume, variety, veracity, and velocity. The volume of data in Big Data analytics is substantial, this requires storage, transfer, processing, and access; technically, these are solved issues, however, established solutions will require implementation (in the form of data platforms) and maintaining data services is a continuous overhead cost. Data variety is crucial to understanding complex systems. It is through the 'multi-dimensional' analysis of physical, biological, and human processes that ecosystem evolution can be interpreted. At the theoretical or causal levels ocean ecosystems are far from fully understood. Therein, variety is doubly important since the most important sources of data (and, in particular, combinations of data) may not be evident *a priori*, or when analysed in isolation. Collecting a sufficient variety of data is challenging in its own right, introducing issues of data interoperability, and adding to issues of data volume, especially when considered over large timespans and with high resolutions. Velocity pertains to (near) real-time applications, and the demand for new data to be accessible with low latency; the challenges associated with this are specific to data collection methods (non-geostationary satellite observation, for example, can have an inherent latency between repeat measurements based on orbital trajectory), data format, pre-processing demands, and technical architecture. Lastly, data veracity is considered in more detail in Section 5.2.3, but for any new source or type of data, or data analytics system, ground truthing and validation is required, and quantification of uncertainty is necessary. Furthermore, the propagation of uncertainty through analysis and into synthesised metrics must be predictable. Issues of trust and veracity are also key when considering the use of AI-based and unsupervised systems.

[15] https://www.go-fair.org/fair-principles/. We also note that NASA has been a proponent of FAIR data principles, in general of Open Source and Open Science data policies. NASA earth data has been free of charge and accessible to researchers, for decades. However, the challenge of maintaining that policy is increasing as data sets grow, and interoperability becomes more of an issue.

Practically, using Big Data and meeting these needs requires data sharing and analytics platforms. At the simplest level, such platforms serve two purposes. The first is to connect end-users with data, translating their needs from the context of their domain to one of data, and intelligently highlighting data that is fit for their purposes. The second – but of equal importance – is to facilitate data sharing. Standardising, normalising, formatting, and sharing data is time consuming and costly. Whilst we may advocate for open data, the reality is this task has costs and associated technical challenges, particularly if the shared data is to have sufficient visibility for use. A good data sharing platform would seek to handle as much of this as possible, also for its own benefit of guaranteeing consistency across data. Architecting an ideal data sharing platform is a task unto itself, and there are numerous organisations either doing this internally, or facilitating data sharing more broadly (e.g., NASA and Global Fishing Watch). The idea of a singular data-sharing platform can be attractive but considering data multi-use across domains ranging from ocean biology to climate science, to economics, and to governance, a multitude of interoperable distributed platforms is likely ideal. To prevent new issues of fragmentation, it is important that such platforms could each access all the openly available data (and avoid per-platform data duplication, which incurs costs and causes issues), but be designed to best support their target sectors.

Northern elephant seal. Photo: NOAA Fisheries

The elephant in the room is how one incentivises data sharing. Traceability plays a key part in this; relating use and benefit to those who provide data inherently increases recognition. For private organisations this allows them to demonstrate their contribution to addressing issues of global importance, such as ecology. For scientists and researchers, it is of even greater importance, as it allows them to demonstrate the benefit of their research – necessary to establish continued funding. One might also consider ways to mandate data sharing. For example, as a condition in grants of funds, or in high seas policy.

There is a significant human element to matters of data sharing, constituting a change in culture. A popular soundbite is that data are the most valuable resource in the world; this encourages the notion that data are a rivalry resource, and that sharing is to one's own competitive detriment. For all but the largest data collectors, this simplistic notion has been shown to be untrue[16]. The effectiveness of data to provide insights in complex domains requires not only data quantity, but substantial variety. This can usually only be realised in the combination of data from numerous sources; data are generally non-rivalry, and sharing is to the mutual benefit of the sector. Establishing this shift in culture is a challenge and is an issue of trust to share. It may be the case that this is best catalysed by independent and trusted organisations that can act as data masters, facilitating sharing whilst protecting the confidentiality of data partners – whether private organisations or nation states.

We may summarise these points into some top-level technical and non-technical requirements for Big Data for Ocean Governance, which may help solution engineering by the development community.

Technical requirements:

- Understanding ecosystems requires Big Data constituting earth observation, surface, and sub-surface data sources, with high volumes of data over long timespans, often high-resolution, and of varied modalities. Therein, free, and open data sharing is a priority need.
- This requires the standardisation and integration of data across communities, particularly datasets from the ocean physics, geo/biochemistry, ocean biology, and human activity monitoring communities. At a minimum, the need for standardisation includes: meta-data, formats, protocols, metrics of certainty, temporal information, depth regimes, and ocean spatialisation/gridding. Adopting FAIR data principles may be an effective path towards this.
- Data sharing platforms are needed to facilitate both data access and data sharing; reducing barriers at both ends. Ideally data sharing platforms should help end-users understand what data are right for their needs, translating domain-expertise to data requirements. Similarly, they should handle as much of the data assimilation process as possible, reducing the cost and challenge of data sharing, and improving consistency across datasets.
- Due to the different needs across ocean sectors, and the multi-disciplinary use of data, a distributed but interoperable model of data platforms is likely needed. Data replication should be minimised to avoid excess costs and errors from duplication, necessitating common access protocols and easy searchability.
- Data access and latency requirements differ based on intended use. Ocean ecosystem analysis is likely to require long-term data of many modalities. Conversely, response and enforcement may require (near) real-time data and analytics. This places requirements on data sharing architectures, as well as on data-collection.
- We need to better understand what 'good' looks like for ocean data, especially in terms of granularity/resolution. Both nuanced governance and ecological analysis may require data of a greater temporal and spatial resolution than conventionally captured; a baseline for analytics needs to be established through testing. Furthermore, significant quantities of data already exist, and are being actively collected; we must analyse where value can be added.

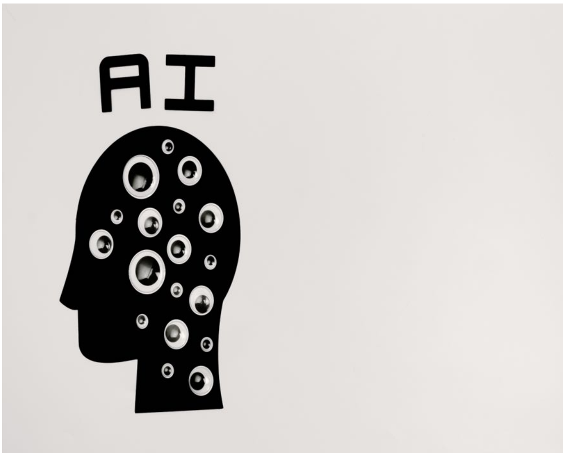
[16] Jones, Charles I., and Christopher Tonetti. "Nonrivalry and the Economics of Data." American Economic Review 110, no. 9 (2020): 2819-58.

Human requirements:

- The foremost requirement for any Big Data solution is to be interpretable to the end-user community. It must reduce barriers and costs to technology access, enable ubiquity, and help the end-user access the 'right' data for their needs. For Ocean Governance, this means that Big Data systems must provide actionable information and knowledge to ocean managers.
- Establishing free and open data sharing is as much a technical challenge as it is a human one. Data sharing must be encouraged and incentivised, there are several ways this could be achieved:
  - Specific positive recognition of those who share data. Particularly towards helping scientists, researchers, and NGOs demonstrate the benefit of their data gathering activities to secure future funding.
  - Establishing methods of trusted data sharing, involving respected, neutral, organisations as data masters, reducing the perception that data sharing may reduce individual competitive advantage, or cause risks to security.
  - Better communicating the need for diverse data, and the non-rivalry nature of data.
  - Mandating data sharing as an output of research grants and introducing policy towards high seas data sharing.
- The cost of sharing and managing data should not be ignored, and focus should be placed equally on facilitating data sharing and data access.

5.2.2 Artificial Intelligence for Ocean Governance

Computing power, quantity of available data, and data gathering have historically been barriers to both Big Data analysis methods and Artificial Intelligence based approaches. It is only relatively recently, with modern computational capabilities and data infrastructures, that these approaches have become viable. They present a natural evolutionary step in data analysis, modelling, and prediction – and as is already the case in many sectors, it is likely that they will come to underpin the methods used for ocean ecosystems analysis, and therein evidencing and implementing Ocean Governance.

Photos: Tara Winstead (Pexels)

Artificial Intelligence is a broad term, covering a wide range of methods, algorithms, and implementations, each with varying capability and maturity. At the simplest level, AI methods can be used to automate statistical data analysis processes, intelligently responding to context based on pre-defined rules, and applying logically straightforward (but perhaps algorithmically advanced) strategies to self-improve analytic accuracy or performance – typically by comparing predictions to reality, and optimising based on tuneable parameters. This can provide unique advantage in terms of speed and scalability, necessary for analysing oceans of data as opposed to constrained test-cases, but such implementations tend not to provide unfamiliar capabilities.

At the simplest level, using AI in this way, typically for pattern recognition, pre-disposes that we know what patterns to look for, and in which data to find them. The weakness of rules-based approaches is that they demand *a priori* understanding. For human observation tasks this can be sufficient, especially in flagging the 'negative case' – unexpected behaviour. This has use from detection of potential illegal activity to emergency response and rescue. Furthermore, the more policy specifies or constrains accepted behaviour, the easier it becomes to automatically identify deviations from it. A concrete example would be using AI to recognise if vessels deviate from shipping lanes/expected routes, or if they enter MPAs.

One might fairly ask, "what about that is 'intelligent'?" The answer: potentially various aspects, each to varying degrees. The superficial task of applying rules to information can fall under the umbrella of AI, particularly when it has ways to self-improve or optimise. Logically, this can be seen as a way of capturing knowledge and expertise in algorithmic form, towards automation and scalability. In our example this might be codifying the knowledge applied, and tests made, by a human expert operator to understand if a vessel's activities are potentially illegal. This is the most rudimentary form of AI, but it is powerful, allowing analysis (of Big Data) to be scaled up beyond what is practical for human operators, allowing a greater variety of data to be considered where analysis is complicated, and can improve technical equity – presenting a cost-effective path for small and developing nations to harness the forefront of expertise. Therein, rules-based approaches are mainly used to assemble unambiguous data, based on *a priori* understanding, and automatically translate that data to knowledge, tailored to decision making or response.

Unfortunately, data are not necessarily unambiguous or machine interpretable. Many aspects of data pre-processing, normalisation, and standardisation can be automated through rules-based approaches, for which there are mature methods in data-driven technology sectors, needing only translation to the ocean domain. However, machine-interpretability is not so straightforward. A pillar of recent AI development has been image and video recognition and enhancement; making evidence and footage that is naturally interpretable to humans also interpretable to machines and using AI techniques to reconstruct 'clear' images from ones that are somehow obscure[17]. This can be generalised to data enhancement and analysis; to an algorithm an image of a human face is not deeply different to that of a vessel, or a weather formation, nor from an 'image' of an ocean ecosystem (perhaps not in colours and intensities of light, but some other coherent, spatially spanning, sets of data).

[17] From resolution enhancement, to denoising, optical image correction, colourisation (e.g., of historical black and white footage), to other forms of reconstruction – all of which are utilising a complex deep knowledge of how 'real' images can look to enhance and repair source images.

Returning to our example, before rules can be applied to check if a vessel's behaviour is illegal, an AI system would first need to identify a vessel, its type, location, and trajectory, based on e.g., satellite imagery. AI capable of this is substantially more advanced than rules-based methods, and truly self-learning. They are built on neural nets and deep learning approaches, and to be effective must be trained by providing the self-learning model many images that have been accurately classified already[18], and allowing it to iterate until it can automatically, successfully, identify and reconstruct features. There are several examples of research into image analysis-based vessel identification and behaviour prediction[19]. AI image recognition is also being used to actively monitor fish catch, being able to identify species and length in real-time through a stereo camera configuration on the trawl rig of ships, with a view to reducing bycatch[20] (and very possibly contributing to enforcement). Furthermore, AI reconstruction tools are also being developed to improve ocean surface temperature data extraction from earth observation[21]; improving data retrieval through cloud cover, and with the further potential to retrospectively enhance datasets. So far, this discussion has centred around using AI to perform analysis that humans are capable of, through both straightforward and advanced methods. Before we go further, it should also be noted that there are many established methods in ocean analysis, such as statistical approaches for fisheries stock assessment, that have proven themselves sufficient. The impulse to use new methods for their own sake (or, similarly, to generate data for its own sake) should be resisted; AI (and any other analytic method enabled by Big Data) is an addition to the existing set of tools, neither an implicit replacement for current methods, nor a 'silver bullet' for ecosystem analysis. With that in mind, let us discuss how AI approaches can move our knowledge and expertise forward, rather than relying on it. The use of AI becomes more interesting when used to analyse complex, highly correlated, systems. With the mixture of climatic, physical, biological, and human properties, oceans and sea-basins almost certainly exhibit complexity. This is more than a descriptive term, it specifically implies that an (eco)system must be described as a whole, exhibiting emergent properties that are not evident from considering each part and process in isolation. This makes analysis challenging for two reasons: firstly, causal relationships are not necessarily evident, with many small factors contributing to an outcome of scale, but no singular, dominant, cause; secondly, complex systems can be unstable, making predictive analysis, and scenario analysis, challenging.

---

[18] Developing a sufficient training set is a great challenge in its own right. A significant approach to this has been through crowdsourcing image identification, whether through quasi-mandatory approaches such as Google's Captcha, or through citizen science and positive engagement, such as the Zooniverse project (https://www.zooniverse.org/).

[19] Literature on this topic is growing, but for some contemporary examples see:
Verbancsics, Phillip, and Josh Harguess. "Image classification using generative neuro evolution for deep learning." In 2015 IEEE winter conference on applications of computer vision, pp. 488-493. IEEE, 2015.
Wang, Senjie, and Zhengwei He. "A prediction model of vessel trajectory based on generative adversarial network." The Journal of Navigation (2021): 1-11.
Guo, Weiya, Xuezhi Xia, and Wang Xiaofei. "A remote sensing ship recognition method based on dynamic probability generative model." Expert systems with applications 41, no. 14 (2014): 6446-6458.
Li, Dan, Hang Liu, and See-Kiong Ng. "VC-GAN: Classifying Vessel Types by Maritime Trajectories using Generative Adversarial Networks." In 2020 IEEE 32nd International Conference on Tools with Artificial Intelligence (ICTAI), pp. 923-928. IEEE, 2020.

[20] For example, see the experiments of Garcia, et al (2020): Garcia, Rafael, Ricard Prados, Josep Quintana, Alexander Tempelaar, Nuno Gracias, Shale Rosen, Håvard Vågstøl, and Kristoffer Løvall. "Automatic segmentation of fish using deep learning with application to fish size measurement." ICES Journal of Marine Science 77, no. 4 (2020): 1354-1366.

[21] A complete review of application of AI methods to earth observation far exceeds the scope of this report. However, Lary, et al (2018) provides numerous good examples of what has been achieved; although the paper is not focused on ocean science, the specific example of characterising pelagic habitats within coastal waters is provided. See: Lary, David J., Gebreab K. Zewdie, Xun Liu, Daji Wu, Estelle Levetin, Rebecca J. Allee, Nabin Malakar et al. "Machine learning applications for earth observation." Earth observation open science and innovation 165 (2018).

In highly controlled, man-made, domains (such as semiconductor and aerospace design) we can capture and fully model complexity within a design process. For global and natural phenomena, we need alternative approaches capable of interpreting data of scale and variety, to identify patterns and relationships that are not pre-defined or known. This is serviced by AI, particularly by Deep Learning and Generative AI methods. These require large quantities of data, often of the greatest variety possible, and over large timespans. With these varied Big Data, such AI methods attempt to reconstruct what they 'see'; seeking to infer a model – based on the correlations across all data sources – that describes the behaviours and trends in the observed system whole. Whilst these approaches can be given starting knowledge to accelerate learning, fundamentally they build their predictive capability independently – not requiring *a priori* understanding of the system from their designers or operators. This can lead to nearly unbiased answers as to what governs ecosystem behaviour, how it is likely to evolve, how specific scenarios would impact it (for example, the impact of a hypothetical offshore wind installation on its surrounding ecosystem), and what sources of data are most important (knowing which is crucial to constraining the scale of Big Data).

Photo: Tara Winstead (Pexels)

Inherently, these approaches have a weakness in that they are predicated on data completeness; no method can understand something that it does not see. This is also a fundamental point at which human bias may be inserted, whereby an AI can be partially blinded through cost constraints, misassumptions, errors, or intent. Designing AI that can understand if it has the right quality and types of data is an area of active research, but it is challenging, and implementations are application specific[22]. It is usually through use, test, and iterative system design that such issues are resolved.

[22] There are numerous approaches to this problem, including ancillary overseer systems that model the performance of the AI system, and self-observation methods; the issue is central to AI system validation, and is an area of very active research. At a very high level, approaches tend to focus on statistical analysis of correlations between inputs and outputs, solution stability, repeatable tests, and analysing system uncertainty (which may indicate missing data sources).

In the ocean domain these approaches are already being used in a limited context, particularly for meteorological and climate analysis[23], and for vessel identification and route prediction. In a wider context they have been used analysing smart cities, through to financial markets, and to assisting medical diagnosis, to provide long-term predictions as well as low-latency situational awareness, feature detection, and alerts. Whilst the principles for these AI approaches are general, implementations are specific. Technology translation towards ecosystem analysis and enhanced governance should neither be seen as trivial, necessarily straight forward, nor free. These tools straddle the boundary of research and application and will require both specific expertise and investment. The competitive demand for AI talent should not be underestimated – there is no surfeit of expertise, and it comes at a premium both in the private sector, and in the context of funding and generating high-impact research. Whilst many sectors may feel as if they are waking up to the potential of AI, and perhaps hoping for technology experts to make their case for its use, bleeding edge sectors are already overwhelmingly competing for this expertise and talent. Therefore, it will be incumbent on those seeking improved ecosystem analysis and governance to build strong links to AI innovators, make known their problems seeking solutions, and incentivise development.

Industry investment in AI tools is already substantial, including in the ocean sector. Increasingly automated operations of complex offshore installations necessitate AI capabilities, for operations, monitoring, and more[24]. Much as with data sharing, it may be that through trusted co-operation cost-effective pathways to enhanced ocean governance can be established, facilitated by the knowledge industry has already developed, but not necessarily disseminated.

Lastly, trust in AI solutions, and AI decision making, must be touched upon. Whilst we explore this more fully in the next section, trust in AI has a challenging technical and human dimension. There is a natural reluctance to place trust in non-human systems. To an extent this can be overcome through comprehensive demonstration and testing, but it does raise deep technical questions pertaining to verification and validation of 'black box' systems. At the very least, deployment of AI systems will require fail-safes and fallbacks and will likely take the form of a series of staggered capabilities. AI systems can be prone to bias, particularly more traditional variants that rely on humans to sort the importance of data, and explicitly specify which data to use (a recent example is the issues face recognition systems have with recognising faces with dark skin-tone, due to having been predominantly trained using light skinned image data, and well-lit images). Means to verify and validate the performance of AI systems must be developed alongside them, at both a human and technical level.

[23] A review of which is presented by Ardabili, et al. (2020). Two specific examples of current advances are Rüttgers, et al's work on generative AI based typhoon trajectory prediction, and Schlör, et al's work on using AI to model sea surface temperature variations in the equatorial Pacific.
Ardabili, Sina, Amir Mosavi, Majid Dehghani, and Annamária R. Várkonyi-Kóczy. "Deep learning and machine learning in hydrological processes climate change and earth systems a systematic review." In International Conference on Global Research and Education, pp. 52-62. Springer, Cham, 2019.
Rüttgers, Mario, Sangseung Lee, Soohwan Jeon, and Donghyun You. "Prediction of a typhoon track using a generative adversarial network and satellite images." Scientific reports 9, no. 1 (2019): 1-15.
Schlör, Jakob, and Bedartha Goswami. "A data-driven generative model for sea surface temperature fields in the tropical Pacific." In EGU General Assembly Conference Abstracts, pp. EGU21-12362. 2021.
[24] For example, see Rahmanifard, et al's review of the uses of AI in the petroleum industry.
Rahmanifard, Hamid, and Tatyana Plaksina. "Application of artificial intelligence techniques in the petroleum industry: a review." Artificial Intelligence Review 52, no. 4 (2019): 2295-2318.

5.2.3 Trust in Big Data & AI

One challenge in technology translation – of any sort – is trust in the new solution. This is particularly the case for data technology and AI, and it is an issue with both a technical and human dimension. The quality, or benefit, of a new solution is irrelevant if lack of trust prevents uptake. In purely operational domains, trust can be generated through demonstration (for example, the use of AI in various forms of autonomy), governance is not purely operational though – it has human, economic, and political consequences. As we have already discussed, a barrier to establishing governance is building sufficient evidence; whether simply via data analysis, or through a complex, black-box, AI-driven system, the methodology must be trusted. Understanding and solving issues of trust are a priority, a perceived lack of trust can easily be used as a justification for inaction; building consensus on what trust means, and how it is achieved, should run before and alongside any technical development – not after it.

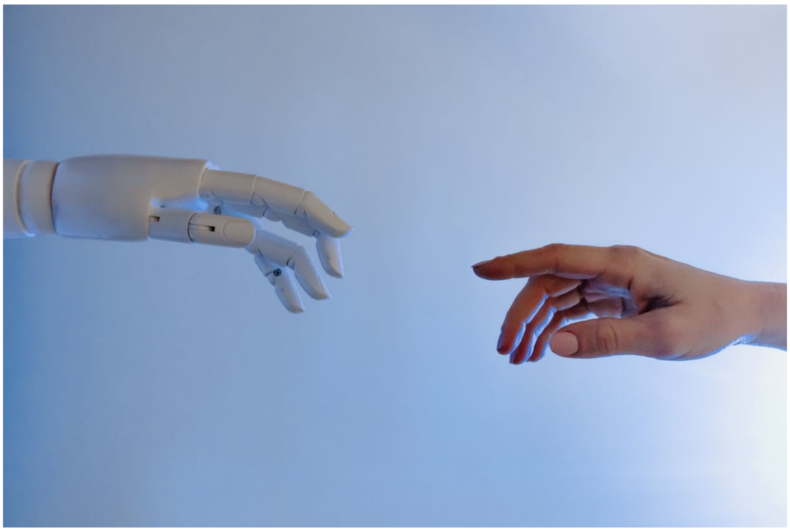


Photo: Tara Winstead (Pexels)

Building trust in data (or any information source) is an old issue, and one that is also well understood. Technically, trust is established through peer review and validation by neutral experts. It involves the scientific community testing and verifying the quality of sensors, sensor deployment regimes, data transfer and collation, and whatever models may sit atop this data layer. This process also involves establishing methods for baselining and communicating data quality, usually in the form of data standardisation and uncertainty metrics. As we move from disparate data sets to Big Data and data fusion, technical standards and data formats will need to be normalised across scientific domains. This is especially true of uncertainty metrics (e.g., range per pixel), as it is vital that data are mutually comparable.

National rivalries are an issue when it comes to data trust, this is best circumvented by data being owned and produced by independent entities. UN driven solutions have an inherent advantage here, and technological solution acceptance needs to be demonstrated by independent bodies such as the UN. This also highlights the importance of collaborative action. People, organisations, and nations have trust in data that they have had a hand in producing and will be more confident in the outputs and analysis of that, whether through institutional or national collaborations, or distributed models of data collection such as platforms of opportunity.

Establishing trust in black-box methods, such as most AI analytics, invokes its own issues. Broadly there are two related problems at the core of this: it is not possible (or practically feasible) to view the workings of black-box systems to manually verify if they are well behaved, and doing what is expected; and, in some cases, the processes within the system are non-deterministic (such as the 'learning' in some, but not all, AI methods). Black-box verification and validation is a domain in its own right, with a multitude of approaches to technical testing and validation. The 'right' tests will depend entirely on what is implemented and how, however broadly, including statistical methods observing both inputs and outputs, verification through use (or on historical data), providing false 'bad' data to verify that bad inputs result in bad outputs, and careful algorithmic design with a view to creating reliable fail-states that unambiguously tell the end-user when the process should not be trusted. It would be an exaggeration to say that verification and validation of AI methods is a solved problem, it is however solvable on a per-implementation basis, and has been addressed in critical, complex, cases such as autonomy.

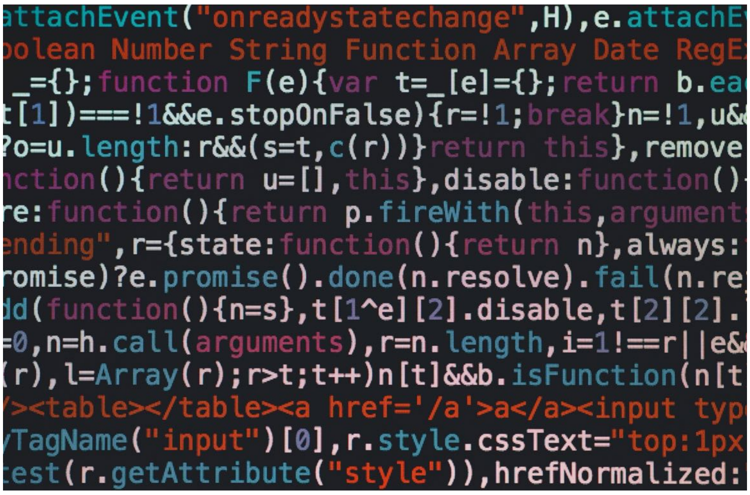


Photo: Markus Spiske (Pexels)

It is important that in developing AI systems these matters of trust and verification are transparently addressed, well documented, and that the wide range of user communities are informed as to why the systems are trustworthy. Whilst it is hard to establish in a competitive environment, open sourcing of tools for public scrutiny should be encouraged. Verification and validation should be continuous, especially for systems that learn, and users should have a way of interpreting the quality of outputs – presumably through a metric of uncertainty, combining data uncertainty with a meta-analysis of the processing. The quality of the data flowing into any system is also paramount; and addressing data issues can be seen as a pre-requisite to the wide use of AI methods. It may help to visually layer AI analytics on top of sensing data, providing end users with an intuitive means to validate what they see. Case studies are also a powerful tool towards building trust and showing success. As we move to AI driven governance, to feasibility tests, to implementations, these case studies need to be built, providing a narrative that contextualises use.

Lastly, as with all novel technologies, fail-safes and fall-back methods must exist. New methods do not invalidate the old, and, at least whilst the technology is nascent, we should work towards a hybrid system-of-systems.

6 Blue Economy Knowledge System (BEKS) Analysis

6.1 Headline Findings

NLAI's Blue Economy Knowledge System (BEKS) is a bespoke market intelligence curation and activation system. BEKS utilises advanced Boolean search string techniques and targeted news alerts to ensure that all relevant market activity is captured and reviewed systematically. A comprehensive, searchable database of over 15,000 filtered Blue Economy (BE) news alerts is already in place, this builds day-by-day. Tailored BEKS campaigns can be aligned to users' needs and areas of interest, to track areas of research, individuals of interest, companies, competitors, territories, and general market conditions. All of these news alerts can be set up to drop into one project-dedicated inbox for daily, weekly or 'as-it-happens' review and analysis. Such intel can feed into market briefings, can help to identify emerging trends, or identify new contacts with whom to engage on our areas of interest.

BEKS draws on all open-source intelligence related to the search terms specified. It works on a standard change detection model, whereby any amendments to indexed web pages are accessed as appropriate. This means that it pulls in all published news alerts, but also changes to academic websites where new papers may be posted (if, for example, a press release about the new paper has not been issued). If any specific academic journals are considered of great importance, these can be specifically targeted.

For this project, BEKS was utilised to capture data relating to the advances of Big Data and Artificial Intelligence in global governance of ocean spaces. This data capture also had the aim of identifying key sub-sectors of ocean governance, i.e., marine monitoring surveillance, marine conservation, and mapping marine ecosystems. By using BEKS, we were able to capture over 140 relevant examples to inform this study. The data ranged geographically, spread across all oceans and seas. The spread of data on AI and Big Data was generous. However, the most common data source was concerned with Marine Monitoring & Surveillance and how Big Data and AI solutions could improve efforts for these technologies.

The use of AI in Marine Monitoring & Surveillance was more popular than Big Data; AI was the more popular term in the search results. This is likely due to the broader public understanding of AI compared to Big Data, which is a lesser-known concept. However, Big Data was a common subject when discussing satellite-based solutions such as remote sensing, earth observation and geospatial data.

Unfortunately, there was not a great deal collected that directly discussed the Sargasso Sea or the potential for technology solutions as a toolkit for ocean governance in this specific region. This may be because there is a lack of awareness or understanding about these solutions and the potential roles they play in ocean governance, or because there has been no practical consideration to expand the use of these tools in the Sargasso Sea. Thus, in this study we conclude that information specific to the Sargasso Sea was more accurately sourced by our Expert and User Group stakeholder engagement.

There are 2 BEKS data "cuts" included as attachments to this Interim Report. The bespoke BEKS inbox related to this project remains "open" and continues to gather useful and relevant information and data. This may be analysed if/as required if further research were commissioned that would continue to build understanding of the key relationships between Ocean Governance & Big Data/AI.

7 Appendices

7.1 Questionnaire responses

A starting point for this project was a questionnaire distributed to the Sargasso Sea Commission's expert and user groups, with a view to understanding the perceived need for Big Data and AI solutions for high seas governance, how consistent this was across different stakeholder groups, and what solutions were already realised. This was not (and was not intended to be) a statistically significant study, with a total of 17 respondents, and a notable bias towards stakeholders from academia, NGOs, and government bodies. Due to these limitations, we do not present a separate analysis of the questionnaire, however, the information gained from it has been reflected in the overall analysis. The questionnaire also provided a snapshot of views, which helped us to structure the interviews discussed in Section 7.2. For the interested reader questionnaire responses are attached to this report.

7.2 Interview responses

A component to this project was targeted interviews with SSC expert and user group members, including representatives of the Sargasso Sea Commission, UNDP, Global Fishing Watch, NASA, and REV Ocean. The purpose of these interviews was to better understand the needs and challenges surrounding ocean governance, both in general, and specific to the Sargasso Sea. Further to this, the interviews explored the potential role of Big Data, artificial intelligence, and remote sensing technologies in enabling and supporting governance and enforcement.

We recorded the interviews in the form of recorded notes; these are not transcripts, and should not be taken as a literal record, free from the interpretive lens of the interviewer. As a general observation, we note that the views expressed by interviewees were in mutual agreement; each stakeholder group elaborated most on their own domain, however their perspectives on ocean governance, technology needs, and even examples of 'good' governance, were in agreement across the board. A caveat to this study is the lack of blue economy/industry stakeholders amongst interviewees, particularly those who would be at the receiving end of regulations, for example in the fishing industry. Whilst one should bear this in mind, we do not feel that it is a limitation of this work; our purpose in this study is to articulate why high seas ocean governance is needed, and what role the aforementioned technologies have to play. We do not seek to comment on the degree, or implementation, of governance that is acceptable, or to recommend policies.

From the interviews ten key themes were identified: The need for governance; Bringing governance to the high seas and the Sargasso Sea; The relationship between policy, governance, and technology; The role of technology in ocean governance; Big Data, and data requirements; The use of artificial intelligence; Remote sensing technologies; Enforcement; Trust in data and AI; and Examples of good governance. These have all been drawn out and analysed in the report body.

7.3 BEKS evidence pack

The evidence from 2 BEKS data "cuts" dated 9th & 24th November 2021, are attached (pdf format for convenience) to this Report.

Project Manager
Kevin Fleming
Kevin.Fleming@nlai.blue

NLA INTERNATIONAL
Registered office: Unit 1 The Cam Centre, Wilbury Way, Hitchin, Hertfordshire SG4 0TW
Registered in England: 10801372