Digital Public

**Part Two of Two**
**Data Sharing Resources**

# Recommendations to improve data sharing agreements for U.S. fisheries in the Pacific region

**Prepared for Intertidal Agency**
**November 2020**

## Table of Contents

This resource guide was written by Sean McDonald, Bianca Wylie, and Vass Bednar of Digital Public. The Net Gains Alliance and Intertidal Agency provided funding for the work, and Kate Wing provided key guidance and support throughout the process. Any errors or omissions in this report are the sole responsibility of the Digital Public team.

# Resources for Negotiating Data Sharing and Rights

## Introduction and Purpose

Part one of this package is a report and analysis of current data sharing issues and opportunities both within NMFS' HMS PSG, and in public regulatory institutions more broadly.[1] That report is a snapshot that identifies and helps prioritize the core issues defining the HMS PSG data ecosystem. We have included a background section below that draws on some of the key issues from the Part One report and how they tie to this set of data sharing resources.

Data sharing is based on relationships. And, when well-communicated, negotiating and growing data sharing relationships can create significant opportunities to increase the value and expand the impact of data. While many of us experience data sharing relationships through formal agreements, like contracts, or specific technologies, both of those things are meant to reflect our relationships - and like relationships, they can be uneven, hard-to-approach, and need attention to adaptation over time. To date, the fishing ecosystem's relationship to data sharing agreements comes, primarily, from compliance requirements, as opposed to a broader relationship development and opportunity space. And for public institutions, ensuring compliance is vital. But data sharing and digital transformation also offer a number of shared opportunities that extend beyond regulation.

The resources that follow are designed to support new approaches to negotiating the benefits, standards, and risks of data sharing relationships. This resources section begins with an overview of some of the key considerations, relationships, ideas, and issues related to data sharing. They are presented to support a general education and level background setting for the things that all parties need to consider when creating a data sharing agreement. They are offered in an order that is not perfectly linear but use the life cycle of a relationship as a way to explain and highlight critical components of negotiating a data sharing agreement. Next, there is a sample and generic data sharing agreement. The purpose of this sample agreement is not to be used as a copy/paste template, but rather to present the outline of an agreement that is marked up to indicate the kinds of issues and conversations that are had to develop the terms of a data sharing agreement. The final piece of this resources section is a data sharing language repository. It was designed as the starting point for a living glossary of the types of terms, phrases, and concepts that are useful to data sharing relationships and agreements. These things change themselves, and so the repository is designed for edits, additions, and growth - it should expand and change over time.

One of the hardest parts of engaging in data sharing and management conversations is that problems typically manifest in tangible, operational ways (uncertainty about whether something can be done, lack of technology to support a proposed approach, etc.) whereas the solutions are often framed in legal processes and terminology. That disconnect is practical and cultural, meaning that it separates the people who experience data sharing problems from many of the people and processes involved in solving them. It is our intent that this resource section bridges those gaps and helps support future data sharing agreements and processes.

Finally, neither this guide nor the Part One report should be construed as legal advice or government guidance. These resources are offered as an educational resource to understand the relationship between data, relationships, and common legal expressions. The authors have

[1] The U.S. National Oceanic and Atmospheric Association (NOAA) is the overarching agency that includes the line office of NOAA Fisheries, or NMFS, and the Highly Migratory Species Professional Specialty Groups (HMS PSG) is a working group within NMFS.

brought their expertise on data policy and governance to bear on U.S. fisheries here, in order to support some of the complex but essential conversations needed for better data sharing and digital transformation. These tools are a starting point for discussions around issues and agreements, and those conversations should engage decision-makers and legal staff prior to finalizing.

**Background from Part One - Data Sharing Summary Report**

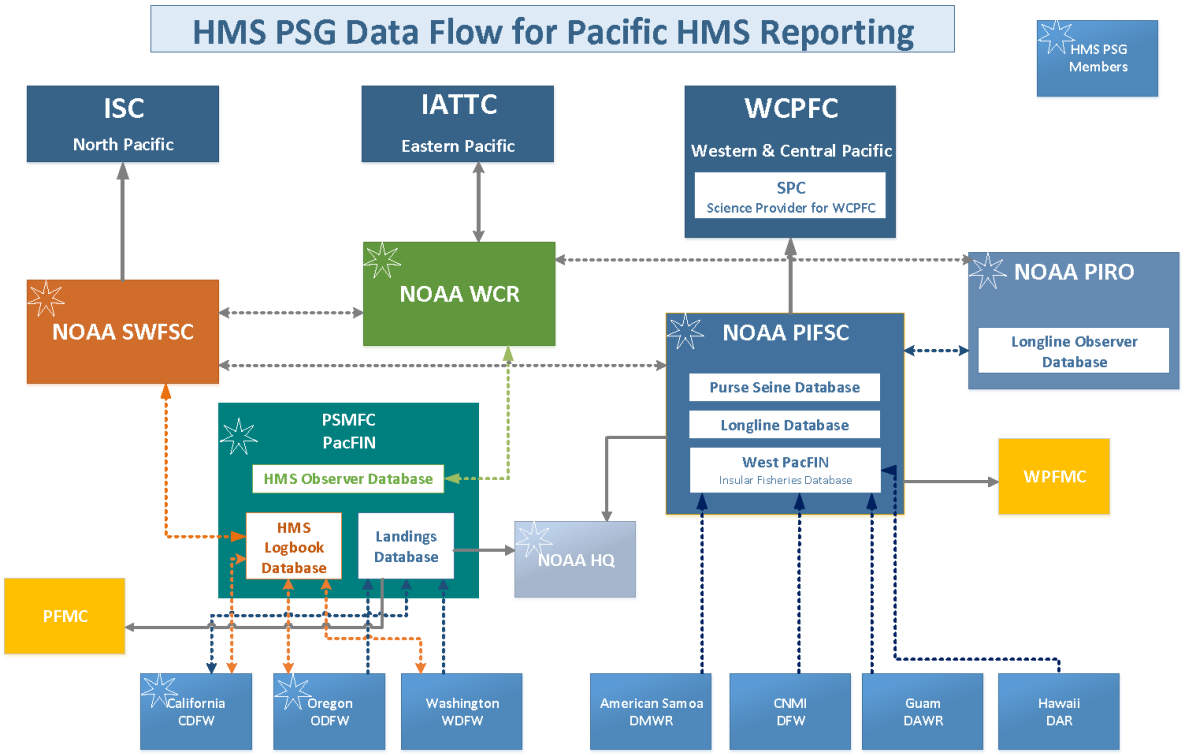
Many HMS PSG stakeholders understand the major trends giving rise to friction and challenges in data sharing, like the need for modernization to meet the goals of the Magnuson Stevens Fishery Conservation and Management Act ("Magnuson"), or clear guidance on how to adhere to the confidentiality requirements of Magnuson while also adhering to open publication requirements of the Open Data Act[2] and the standards set through NOAA's Data Strategy. While those issues are known, they aren't clearly connected to day-to-day problems, or visible in aggregate, meaning that people experiencing data sharing problems often, understandably, view each challenge as minor, standalone, and, sometimes irrational. Even for those involved in managing fisheries data, what is framed as "the problem" is often a *symptom* of a larger, unresolved tension. That disconnect can add frustration and complexity to digital and cultural transitions, which are difficult enough without it.

The fishing ecosystem, to its credit, has a strong informal culture. Data sharing throughout the HMS PSG, until recently, was managed through NOAA rule interpretations, often private memoranda of understanding, and a patchwork of norms and relationships. While that informality facilitates intra-community data sharing, it can also act as a barrier to expanding engagement because these approaches do not have the processes or trusted relationships to adapt to new circumstances. Informal arrangements also adapt poorly to change, because they are often difficult to iterate on - especially when they are nested underneath regulatory reporting requirements. This issue is especially complicated in the data workflows and multi-actor supply chains that are common to the HMS PSG and its research (see diagram below)

[2] S. 760 To expand the Government's use and administration of data to facilitate transparency, effective governance, and innovation, and for other purposes.

Figure One

**Highly Migratory Species Professional Specialty Group Data Flow for Pacific Highly Migratory Species Reporting (provided by HMS PSG)**




ISC: International Scientific Committee for Tuna and Tuna-like Species
IATTC: Inter-American Tropical Tuna Commission
WCPFC: Western and Central Pacific Fisheries Commission
SWFSC: Southwest Fisheries Science Center
WCR: West Coast Region
PIRO: Pacific Islands Regional Office
PIFSC: Pacific Islands Science Center
PFMC: Pacific Fishery Management Council
WPFMC: Western Pacific Fishery Management Council
PSMFC: Pacific States Fisheries Management Commission
PacFIN: Pacific Fisheries Information Network
*Blue boxes on bottom row represent state/territory fish and wildlife agencies*

In order to improve its internal data management and to engage with its growingly digital fishermen and partners, the HMS PSG is wrestling with fundamental questions about data sharing. Not only about *how* to formalize data sharing agreements, but also *how formal* data sharing agreements should be. Technology industry contracts, for example, are specific about the rights conveyed with data, like general re-use or only to provide a particular service – even though the rights they often grant are expansive, open-ended authorities over the data involved. Fisheries data contracting, by contrast, tends to speak specifically to the chains of authority and the purposes of data use, but include ambiguous conditions on the authorities involved (like the balance between 'confidential', 'private', and 'open'). Resolving these tensions is a short-term operational cost, and, as evidenced from the research, inaction also creates costs. Without direct intervention, private actors and markets tend to take the initiative for setting data collection, use, and sharing standards, often in ways that fragment and polarize areas of potential collaboration.

Negotiating complementary data licenses with fishermen and data owners, in addition to the rights conveyed through Magnuson, structurally addresses the issues posed by informality and concerns about regulatory reach in two, fundamental ways: (1) it reverts to simple and direct data rights, creating the opportunity to clarify supply chains and workflows, reducing long-term confusion for all parties involved, and building trust; and (2) the licenses center the negotiation of the data sharing relationship in mutual value, helping build cultural and practical buy-in through equitable negotiation. The HMS PSG, and the broader NOAA community, are likely to benefit from negotiating a range of new data sharing agreements, some of which position the government as a peer or partner, in addition to regulator.

We have structured this section the way that most industries approach relationships. The goal is to explain when, how, and why data relationships formalize. The focus of the resources section is to communicate the key elements of data sharing relationships, how they operate in data sharing agreements, and how their design affects the issues raised by the HMS PSG.

## How Data Sharing Agreements Work

Data sharing agreements describe the terms of a relationship in which two parties exchange digital information. Most data sharing relationships start with a purpose and a particular piece or stream of data in mind.[3] From there, the parties typically establish their expectations for the conditions of that exchange, how those conditions are enforced, and how the parties - should they disagree - resolve any disagreements about whether the terms of the contract are being met. Data sharing agreements often spend as much time defining the monitoring and enforcement relationships between the parties sharing data,[4] as they do the specificities of the data to be shared. As the fishing industry is learning through the experience of video electronic monitoring, the data sharing and management issues that arise out of trying to monitor an activity can be more complicated than the activity itself.[5] And, as a result, the main differences between data

[3] For example, in 2012, Canada and the Russian Federation signed a Memorandum of Understanding to deter illegal, unreported and unregulated fishing by requiring vessels to transmit a specific list of data (owner's name, vessel registration number)in advance of entering the other country's ports.

[4] MOU 225-09-0008, a Memorandum of Understanding between the U.S. Department of Commerce, NOAA, and the U.S. Department of Health and Human Services Food and Drug Administration includes information regarding the monitoring and enforcement relationships between the parties sharing data (2009).

[5] In 2016, Spanish vessels signed an MoU for electronic monitoring in the Cook Islands.

sharing agreements are not the specific terms or types of data, but how the parties define and enforce the terms of their relationship.

Data sharing relationships may refer to sharing something as "open data" - which, in its purest form means that anyone can do anything with the data, once shared.[6] Of course, in cases where data can be used to harm people or business interests, pure "open" data is not ideal, in which case, there are often limits placed on a party's ability to share data, such as if it is 'anonymized'. In this case, the implication is that anonymization of the data set will prevent anyone who might use the data from realizing its potential harms.

There are a number of legal terms in data sharing negotiations - like 'anonymized', 'private', 'confidential', and 'secure' - which are concepts most people have a general sense of, but that encompass specific technical practices, industry standards, and a range of legal standards. One of the challenges in negotiating data sharing relationships is not only being specific about what the terms refer to, but also the standards of treatment necessary to achieve their conditions. Data sharing agreements create power dynamics themselves so it is important to ensure that agreements reflect appropriate limits on the power of each party to change the terms, end the agreement, and ensure a fair resolution to disputes.

When it comes to building clear data sharing agreements, the questions aren't just who, what, and how - they are also how will we know, how can we prevent everything else, and what happens if we do not. Data sharing relationships that fail to set meaningful limitations or consider likely threats are as defined by the impacts of their failure as they are their original intentions. Counterintuitively, data sharing relationships and agreements are easiest to approach as defining a set of limitations. Without limitations, there is no way to assert data rights or prevent abuse.

## Data sharing terms and concepts in context

This section presents an overview of key elements of data sharing agreements, with examples framing them in fisheries and other relevant industries. While some of these concepts may seem basic or self-explanatory, having a common understanding of them across all participants in a data discussion will set the stage for better negotiations and a more effective agreement. For a more detailed list of concepts and language to potentially incorporate into an agreement, see Appendix B: Data Sharing Language Repository Appendix B: Data Sharing Language Repository.

This section is designed to unpack and translate the legal terminology common in data sharing agreements to highlight core aspects of data sharing relationships and link them to their use and reference in the Sample Agreement. The headings refer to the terms typically used in data sharing agreements, and the descriptions explain the significance.

### *Who's Involved? (Parties)*

The parties to an agreement are the people who are bound by its terms. In most contracts, the parties section names everyone involved - typically with identifying information, like address, and role in the agreement.[7] The parties section, where

[6] The Australian government (via the Australian Fisheries Management Authority) has enabled the downloading of raw (unprocessed) data on annual catch and effort for Commonwealth fishers.

[7] A 2019 MOU between NOAA, the Bureau of Ocean Energy Management, and the Responsible Offshore Development Alliance describes each entity entering into the agreement as well as the role of the signatories.

appropriate, may also identify or define the roles occupied by each party, such as "employer/employee,[8]" to evoke professional standards and interpretations.

That sounds obvious, but it can get complicated quickly, largely because data sharing agreements are between (at least) two parties, but they also typically limit the way that those two parties can share data with others. In addition, there are ways that data sharing agreements can, explicitly and implicitly, create data rights for groups who are not party to the agreement. Here are a few of the additional ways that data sharing agreements create and limit data rights:

#### Assignment

Assignment clauses are what enables parties to convey their interest or rights to other parties - either directly, through subcontracting, or indirectly, based on a change in circumstances, like bankruptcy or death. The benefit of flexible assignment clauses is that they enable adaptable data workflows; the downside is whether and how the conditions of the underlying agreement will apply to assigned parties.

#### License

Data is typically shared not only through agreements but also through licenses. Licenses are the specific conditions, uses, and rights under which data is shared between parties and the way they frame data sharing rights and limits can give parties the right to share data with others. In this document, we have tried to be clear when we are talking about licensing *for data* versus *fishing* licenses, but the underlying legal framework is the same for both - a set of agreed on conditions allowing an activity – and in some cases a fishing license may include explicit data sharing terms.

#### Conditional Rights (i.e - Confidentiality)

Conditional rights are limitations on data sharing that are contingent on context. For example, confidentiality clauses typically create strict prohibitions on sharing data as long as it isn't common knowledge, but if the protected information becomes public, then those prohibitions disappear.

### ***Introducing the Relationship (Preamble + Authorities)***

Preambles are optional in most data sharing agreements, but they are common between public institutions and between parties whose right to make an agreement is based on another relationship. Preambles can cite the founding authority of the parties[9], the events

[8] An MOU between the Bureau of Ocean Energy Management and the Bureau of Safety and Environmental Enforcement indicates the roles of each signatory (2014).

[9] The Data Use License Agreement template that has been approved by the USGS Office of Policy and Analysis (OPA) has a preamble that asserts that the "USGS is authorized to collaborate with other Federal agencies, units of State or local government, industrial organizations, private corporations, public and private foundations, and non-profit organizations (including universities) under the Stevenson-Wydler Act.

that led up to the agreement,[10] or the mutual priorities and intentions of the parties.[11] Private data contracts often include a provision in which the parties explicitly commit to having the necessary authorities to make the agreement.

### *Purpose*

The purpose of an agreement is important because it is often the frame for, or a condition of, data sharing authorities, meaning that data shared between the parties is only valid if it's used for the purpose of the agreement. One of the main distinctions in understanding data management is the difference between 'purpose' and 'use'. The purpose of a data sharing relationship is, essentially, an intention or desired outcome - and so the validity and quality of data management and sharing can be judged based whether it contributed to the desired outcome. For example, NOAA's confidential data management and sharing are typically required to fit into the limited purposes covered in Magnuson. By contrast, a 'use' of data is more like an activity or action, which may or may not fall within a purpose - but the reason to include a purpose is to make it easier for courts to understand whether a specific action was intended by the parties.

As mentioned above, some data sharing agreements manage this by articulating the intended purpose with broad language, like 'the management of marine species in federal waters', whereas others focus on comparatively specific goals, like addressing 'bycatch of protected living marine resources.'[12] For the data collected under the regulatory mandate created by Magnuson, every use of that data has to have a clear relationship to a purpose outlined in Magnuson.

### *Scope + Data Description*

There are a range of ways to frame the scope of a data sharing agreement, each of which has an implication on impact and sustainability. Scope can also have a determinative impact on what data can be shared, under which circumstances, and by who.

Data sharing agreements are shaped by both the place and time they are created, as well as whether they focus on the data being shared, the activity being pursued,[13] and/or the

---

[10] The fisheries access deal between the US and Pacific countries describes that the agreement was Initially struck in principle in the wee hours of a weekend morning at the 18th negotiation session in Auckland this June, the 2017 Treaty deal was prepped for signature by officials for the US and FFA members in Nadi over the weekend, with more signing done as the 3rd Pacific Tuna Commission met.

[11] For instance, in May 2007, Canada and Norway signed a Memorandum of Understanding on Fisheries Cooperation to advance collaborative activities to ensure the conservation and sustainable international management of global marine resources.

[12] Both of these mandates were created by Magnuson, but by specifying a purpose, parties reserve the right to direct data they share toward specific uses, while at least theoretically retaining the right to challenge or opt-out of illegitimate and overbroad purpose interpretations.

[13] A 2007 MOU between the Ministry of Agriculture, Fisheries and Food of the Kingdom of Spain and the Department of Fisheries and Oceans of Canada explicitly states two mutually desired activities: Motivated by the desire to increase the existing level of cooperation in the fields of fisheries and marine sciences research between Spain and Canada, notably in respect of fishing activities in the northwest Atlantic; and, Determined to reinforce their bilateral cooperation in fisheries relations as a means to consolidate mutual trust and to conserve and sustainably manage the fisheries resources.

intended impact or outcome of the data sharing relationship.[14] For example, in the United States, data is often regulated by whether or not it is personally identifying, and therefore requires a formal agreement to share; whereas many assume that, if data is not personally identifying, then it may not require a formal agreement to be shared. That assumption creates challenges if parties want to share data internationally or to combine the data with other data that would then make it personally identifying. Here, the point is that an agreement that references "personally identifying data" bases the scope of the agreement on the characteristics of the data.

By contrast, data sharing agreements can frame the scope around provenance (source), or supply chain of data collection. In this case, as long as the data is collected and transferred legally, then the parties are free to use the data however they wish. While this type of frame ensures continuity of rights in important ways, it can also be abused based on the contextual nature of data's value and risks. Another common approach to defining the scope of data sharing agreements is based on the common purpose, mandate, and/or role of the agreeing parties - usually with language like "parties agree to share all available data in order to rebuild overfished stocks of highly migratory species in the Pacific." In this approach, the volume and type of data shared is determined almost exclusively by its perceived utility to achieving the particular goal of rebuilding overfished stocks of migratory species in the Pacific.

These approaches are not mutually exclusive and, in fact, can be used as layers so that parties, for example, are only able to share data collected in particular ways if it is useful to a shared purpose. Each approach to defining the frame of the data collected highlights the priorities, scope, and clarity of the parties involved and the use they envision.

### *Explaining the Data and its Use (Defining Data)*

Data is not any one thing and the law that applies to data typically varies based on the context of its collection, contents, use, and transfer. There are a number of legal approaches to defining data, whether as property, as personally identifying information, and as a representation - to name a few. As a result, data sharing agreements aimed at formalizing a particular data workflow or type of data relationship create rights based on that context as opposed to the entirety of ways the data could be used.[15]

This same approach is why data, under Magnuson, can be treated as a trade secret. At the time the statute was written, a primary purpose for that information was to understand, mitigate, and equitably apportion the effects of commercial fishing. While that remains one goal of the law today, the mandates in the law have expanded to include multiple sectors and the broader ocean ecosystem. As a result, the data collected under Magnuson's mandate is expected to be used by more and different actors and, in different circumstances the same data element can play the roles of trade secret, public record, an affirmative representation for law enforcement, work product delivered through

[14] The intended outcome of the United Nations Food and Agriculture Organization (FAO) agreement – the Port State Measures Agreement (PSMA) – is to provide the means for governments to require signatories to collect and share data on fishing vessels entering their ports and to deny entry to vessels that have been fishing illegally.

[15] For instance, the National Science Foundation defines "data" as information relevant to, or of interest to social, behavioral, and economic researchers, either as inputs into or outputs from research. Data includes research materials resulting from primary data collection or creation, or derived from existing sources.

consulting services, personally identifiable information, exhaust from digital platform use, and openly licensed science, among many others.

The most important value to seek in data definitions in sharing agreements is clarity. Because data is cheap and easy to copy, share, transfer, and make derivative products out of, most agreements grant expansive rights to data. It is common to see terms like "global," "perpetual," and "royalty-free," in data agreements, allowing parties to use shared data however they please. That unlimited use makes it very difficult to create contextually relevant rights protections, which increases data use but decreases the amount of control data owners can exert on how their information is used, even if it's against them. A 2019 report from the Duke Nicholas Institute and the Internet of Water, "Assessment of Federal and State Agreements with Data Organizations," (the "Duke Report")[16] does an excellent job of explaining the diversity of contexts and treatments of data in different agreements. To that end, here are the types of provisions to focus on for approaches to, and definitions of, data:

### *Non-Disclosure* + *Confidentiality*

In most relationships, the parties exchange private information - plans, communication about and observation of events, or commercial secrets. Provisions can be drafted to apply to data shared during a relationship, which typically enables the parties to share data internally as necessary to fulfill the main objective, like eligibility for a fishing license, with the expectation that the data is not used outside the scope of that relationship, and is deleted or returned at the end of the relationship.

### *Intellectual Property*

In relationships where data is intellectual property (IP), the focus is typically on three core areas: (1) the ability of each party to use and create value from the data, and how, if at all, that value should be distributed; (2) the rights of the parties to prevent each other from realizing the value (and, in a much smaller number of cases, avoiding the risk) of data use and sharing; and (3) the act of publication, sharing, or granting additional license to the data as the core model of use.

There are too many available models of IP agreements to try and draw best practices beyond the benefits of clearly articulated, bounded data uses and rights - as opposed to open-ended licenses which, initially, enable actors to exert their interests unfettered, but rarely help stakeholders or parties resolve disputes.

### *Applicable Law (Reporting)*

Beyond fisheries, specifically, there are a number of other adjacent regulations with an effect on data, such as business reporting requirements or data retention. It's important to understand those laws, and how they may constrain the parties' ability to make guarantees, especially about independence from law enforcement.

[16] "Assessment of Federal and State Agreements with Data Organizations", Duke Nicholas Institute for Environmental Policy Solutions, Internet of Water. November 2019.

### *How Long Should this Go on For? (Term)*

The term of an agreement is how long the parties expect to be involved in the data sharing relationship. Thankfully, term provisions are simple.[17]

The key distinction to observe in negotiating the term of data sharing agreements is *the term of the agreement*, as opposed to *the term of the rights conferred through the agreement*. Many data sharing agreements, such as those compelled by regulators, don't focus on a single data point or database, but rather create ongoing data transfer as long as the agreement is in place. So, the term of the agreement itself bounds the amount of time that the parties share access to data - as opposed to limiting how long the parties can use the data shared under the agreement. This is in contrast to data-specific licenses and limitations, which determine how long, and under what conditions, the parties can share data.[18] The term of the agreement is the time during which each party collects and shares data.

### *Describing the Standards, Conditions, and the Limits (Data Management + Governance)*

As digital transformation continues apace, so do the number of professional standards and approaches to nearly every aspect of the data life cycle. One of the main reasons that data sharing agreements are so important is that it is extremely hard to build perfect security or un-hackable infrastructure. Rather than expect perfection, an increasing number of data sharing agreements rely on parties to take 'reasonable' measures to collect, manage, use, transfer, store, and delete data, among other things.[19] While these activities are supported by a growing range of technical, legal, and professional standards, they still create the unavoidable potential for risk.

One of the main functions of data sharing agreements is to clarify and apportion the risks, or potential for risks, posed by data transfers. Most data sharing agreements vary based on the degree of liability the parties take, how mutually or equally they share the blame for breach, and what kinds of assurances or steps the parties are able to make to secure their systems.[20]

---

[17] Of note, a 2005 MOU on Fisheries Cooperation between the Department of Fisheries and Oceans of Canada and the Ministry of Agriculture, Rural Development and Fisheries of Portugal indicates that it will remain in effect for an "indeterminate time."

[18] A 2014 MOU between the U.S. Department of Defense and the U.S. Fish and Wildlife Service to promote the conservation of migratory birds specifies that the Fish and Wildlife Service must "Provide essential background information to DoD, when requested, to ensure sound management decisions. This may include information on migratory bird distributions, status, key habitats, conservation guidelines, and risk factors within each BCR."

[19] The National Geospatial Advisory Committee published a primer on Interagency Data Sharing that recognizes the challenges of security, privacy protection, and licensing, among other factors.

[20] In contrast, a Memorandum of Understanding between the National Energy Board and Fisheries and Oceans Canada for Cooperation and Administration of the Fisheries Act and the Specifies at Risk Act Related to Regulating Energy Infrastructure explicitly addresses legal liability, clarifying that the "MOU is an expression of mutual intentions of the Parties and is not legally binding on them or enforceable against them (2013).

In most professional and commercial contracting, standards of professional conduct are an important way to ensure that each party takes reasonable care in the tasks they perform. One of the trends in data sharing agreements, and technology contracting more generally, has been a move away from most forms of liability, including professional standards of conduct, as well as basic guarantees about how the data or project works. Whereas many industries make specific guarantees or warranties about their products, often with the caveat of the conditions and maintenance required for their use, many commercial data sharing agreements broadly disclaim the quality and accuracy of the data, in addition to broadly indemnifying each other. The removal of those traditional liabilities can also reduce incentives to build or articulate professional standards for data management.

The shift in liabilities fundamentally affects the role that data sharing agreements play. Historically, these agreements conveyed a limited amount of discretion and trust, which was secured by the other party's willingness to assume liability for abusing that trust. The move away from liability also removes the guarantees about how each party manages or uses data - and as a result, data sharing agreements have changed in two overarching ways: (1) they are increasingly specific about the what, how, why, and when each party is able to share data between each other, to control external liability; and (2) they are increasingly transactional - assuming that the parties will not interact and removing the provisions that enable parties to adapt, challenge, or dispute elements of the agreement. Data sharing agreements are not alone in managing the tension between transactional contracting, which governs temporary interactions, and relational contracting, which supports longer-term relationships. As described in the 'term' section - this tension is complicated by the difference between the length of time the parties are afforded the right to use data (the license) and the control the parties wish to exert through the relationship (the contract).

The effect of those shifts, for fisheries data sharing agreements, is that the parties involved have the opportunity to explicitly define data sharing as a transactional or longer-term relationship.[21] The questions and tensions described above have largely arisen through a pattern of behaviors and incentives, rather than a clear or intentional set of decisions by a formal group. One of the opportunities that negotiating new data sharing agreements offers the HMS PSG and NOAA is to shape the incentives that will guide the liability framework for future fisheries data management.

When designing these agreements, there are three, foundational questions the parties need to answer: (1) How formalized does the data life cycle need to be? And by life cycle, we mean the workflows that guide the various stages of data use – collection, storage, archive etc. – rather than any specific definition of the term; (2) What standards and/or behaviors should the parties adopt?; and (3) How do the parties amend and/or adapt the agreement?

[21] The University of Florida and the George A. Smathers Libraries have an example draft data transfer agreement that encourages parties to clearly draft the term of the relationship: Data Transfer Agreement: Statewide Ecosystem Assessment of Coastal and Aquatic Resources (SEACAR) [PDF].

*(1) How formalized does the data life cycle need to be?*

The more people and organizations depend on the output of a data workflow, the more important it is for that relationship to be formalized. Similarly, the more ongoing a data relationship, the more important it is to formalize key elements of the relationship - like how parties might replace a data standard or adapt to include other parties. Some data sharing agreements, particularly those that govern highly regulated data sets, like health records or minors' data, are very specific about data management practice[22] as a way to manage liability. Other data sharing agreements focus on ensuring that the data transfer is valid, but otherwise grant broad licenses and disclaim any liability that might arise from the data or the sharing. The parties to a data sharing agreement first have to articulate what kind of relationship they have, and how much of it they want to formalize in an agreement.[23]

There are two main elements to consider articulating in data sharing agreements: (1) what technical standards and expectations does each party have for the relationship; and (2) how will the parties manage unanticipated problems together - and/or against each other? The degree to which workflows are formalized will help make these conversations easier to have, but they aren't a necessity and it's common that an organization has a mix of formal and informal workflows in their processes. It is not a condition to have them all as formal as possible prior to engaging in new data sharing arrangements - simply being aware of this factor of formality and informality is helpful to these conversations. Legible and formal workflows help identify who to engage with and how to engage with them, and for this reason increased formality is helpful from an operational and efficiency perspective.

(2) *What standards and/or behaviors do the parties need to adopt?*

The two primary things that a data sharing agreement formalizes are how the parties manage data and how the parties agree to handle disputes and future challenges. In the case of the former, there are a range of competing legal and technical standards for defining the "best" approach to data management challenges. Standards are rules of practice that are often independently certified trade organizations or public institutions. In the United States, the National Institute of Science and Technology, for example, maintains standards for data integrity, security, and sharing. In addition, there are large, international technical standards federations, like the International Organization of Standardization (ISO) and the Institute of Electrical and Electronics Engineers (IEEE), both of which generate standards for how to manage and share data. Both bodies are known for their rigor so setting standards is less about picking the 'right' standards, than about how to

[22] The Health Insurance Portability and Accountability Act (HIPAA) sets the standard for sensitive patient data collection. Companies that deal with protected health information (PHI) must have physical, network, and process security measures in place and follow them to ensure compliance.

[23] A recent reconciliation agreement between the Government of Canada and B.C. Coastal First Nations seeks to broadly improve fisheries governance – part of this agreement focuses on creating increased clarity about the roles and processes related to information sharing.

effectively monitor and enforce those standards. Certification bodies are one way, but they do more to ensure general practice than prevent specific harms.

Beyond ensuring data management professionalism, data sharing agreements typically include assurances as to the quality or fitness of a data or technology product for the articulated purpose. Warranties are the guarantees that a party makes about the quality and necessary conditions for use of its products. By articulating a guarantee, the party assumes legal liability in case the product does not meet those standards and, by implication, disclaims liability for things that fall outside of those uses. While traditional commercial contracting often features warranties as a sign of quality, digital and data contracting often make no guarantees about data quality or, even, legality of collection. One of the reasons for this is that data can be used for many things, and so warranties that limit future use are seen as anathema to the technology industry's focus on unbounded innovation. Warranties, indemnity clauses, and disclaimers are all important ways that data sharing agreements frame and, more commonly, limit the liabilities that parties assume for the integrity of the data sharing ecosystem.

In data sharing agreements, standards and warranties are a way to set expectations, and to articulate the consequences of failing to meet them. When misused, standards can be a way to generalize away liability for failing to prevent predictable harms - like poor due diligence or oversight. Data standards are a way to avoid being deterministic about how each party manages their work, while upholding quality.

*(3) How do the parties amend and/or adapt the agreement?*

This is typically managed through the amendment authority in data sharing agreements, which sets out the conditions for changing the rules. The ability to change the rules is as important as the initial agreement, as whoever can change the rules can subsequently alter its terms. The primary difference between amendment provisions comes down to how mutual the authorities are, and how granular the changes are. For example, a number of technology platform contracts give the technology provider the unilateral authority to change the terms of service, whereas more directly negotiated data sharing agreements require mutual, written agreement to amend the terms. One way to manage the potential abuse of those authorities is to limit the scope of the amendment provision to minor technical changes, or to ensure that parties do so 'in good faith' and with ample notice to the other party. Ultimately, amendment authority is one of the most important elements of bridging the gap between a single instance of data sharing as an activity, and the broader impacts it has on, and ways it reflects, the underlying relationship between the parties.

The other major unilateral change that most data sharing agreements include is a provision for termination. When termination happens in the normal course of a data sharing relationship, it's either because of the expiration of the term, or the parties have fulfilled the requirements of the termination provision. The primary difference between termination provisions, like amendment authorities, are how equally or mutually agreed upon they're required to be. Most data sharing agreements also

enable ‘termination for cause’ which gives a party the right to unilaterally end the agreement and/or the rights conferred through it, based on a breach of the agreement. Termination for cause provisions vary based on how much notification the parties need to give and whether there are opportunities to fix the problem in good-faith.

### *Establishing Responsibilities and Solving Problems (Liabilities + Disputes)*

One of the reasons law and the negotiation of legal agreements can be unpleasant is that they require considering how and why things might go wrong. Worse, they often start to assign hypothetical blame before the relationship has even begun. And yet, as technology investors have been known to point out, systems that fail to consider threats are more often characterized by them then by their values. In other words, unless the parties to a data sharing relationship consider how their work could be exploited, their relationship can do more harm than good. Negotiating legal agreements often goes one step further - not only do they require understanding threats, but they also often assign responsibilities for addressing them and the punishments for failing to do so.

Liability is how legal agreements acknowledge different kinds of harm and organize the relationship - or at least offer restitution - based on that harm. Data sharing agreements frame liability, like intellectual property, in context, and so it varies significantly. Two important areas to explore in discussing liability are: (1) the type and scale of potential harm and what responses would address the issue; and (2) how, where, and under whose rules the parties resolve disputes. One challenge with designing liabilities is ensuring they are proportionate to the scope of the relationship and the impact of available harms. In cases where the harm of sharing data significantly exceeds the value of the contract (for example, a trade secret) an agreement may refer disputes to authorities that can levy appropriate punishment. Typically, parties’ liability inside of a contract is limited to financial restitution, based on the parties acting in good faith - and anything that exceeds that, or in cases where either party disagrees that there has been a breach - is resolved as a dispute.

### *What’s at Stake? (Remedy)*

*Penalty + Restitution*

In addition to agreeing what the parties should do, data sharing relationship negotiations often consider how serious the consequences should be if they don’t do the things agreed upon. While this is typically the role of a dispute resolution system, parties can indicate the general size and type of harms that might arise from breach, which can help courts understand the original intent of the parties. This also means that penalties are typically limited to the kinds of things that courts can compel, which are: (a) financial - imposing a cost or fine for a particular action; (b) injunctive - a much more rare, but rule-based prevention on a specific action; and (c) freedom - typically limited to criminal matters, courts can jail and, in very limited circumstances, restrict specific freedoms.

### *Who Solves Problems We Can't? (Choice of Law + Dispute Resolution)*

Data sharing agreements typically point to the institutional and external authorities the parties will defer to for dispute resolution. This usually happens on two levels: the choice of law clause and, often, a dispute resolution clause (sometimes referred to as an arbitration agreement). The choice of law clause affirmatively picks a legal jurisdiction for the agreement, which means that jurisdiction's laws govern the relationships and, if the parties disagree, they will go to court in that jurisdiction. Choice of law clauses are important because they import the applicable laws of that jurisdiction and, essentially, subject the parties to the influence of its public institutions.

Dispute resolution clauses are what govern everything leading up to the decision to go to court. There are a range of dispute resolution approaches, ranging from the parties giving each other good-faith notice and an opportunity to respond to forcing parties into binding arbitration. Dispute resolution clauses, when well-designed, clearly articulate a process for identifying and managing disagreements collaboratively and an escalation pathway that preserves as much of the value of the relationship as possible. When parties decide that they are unable to resolve a dispute themselves, they have the choice of whether to go directly to court, or to work with alternative dispute resolution services - like mediation and arbitration. In these cases, parties typically negotiate and enter into a second contract or agree to an arbitrator's ruling on how they will resolve the dispute. While there are a range of contexts and decisions that go into resolving disputes coming from data sharing agreements, it rarely serves the best interest of either party to forego their right to use the courts.

In some cases, courts will also strike down a specific provision of an agreement. Many agreements contain a provision for "Severability" which essentially ensures that if one provision is found 'invalid' or otherwise illegitimate, it doesn't impact the rest of the agreement.

## Appendix A: Annotated Model Agreement

In the annotated model agreement, we use the following eight symbols/icons to represent the related resources to consider when negotiating data sharing and rights.

 WHO: Who's Involved? ("Parties")

 RELATIONSHIP: Introducing the Relationship ("Preamble + Authorities")

 DEFINITIONS: Explaining the Data and its Use ("Defining Data")

 DURATION: How Long Should This Go on For? ("Term/Period of Agreement")

 MANAGEMENT: Describing the Standards, Conditions, and the Limits ("Data Management + Governance")

 RESPONSIBILITIES: Establishing Responsibilities and Solving Problems ("Liabilities + Disputes")

 REMEDY: What's at Stake ("Remedy")

 DISPUTE RESOLUTION Who Solves Problems We Can't? ("Choice of Law + Dispute Resolution").

# [SAMPLE] DATA SHARING AGREEMENT

## 1. Parties

The agreement between **[PARTY ONE]** located at (address) and **[PARTY TWO/PUBLIC INSTITUTION/NOAA FISHERIES or NMFS]** located at (address).

The parties to an agreement are the people who are bound by its terms. In most contracts, the parties section names everyone involved - typically with identifying information, like address, and role in the agreement. The parties section, where appropriate, may also identify or define the roles occupied by each party, such as "employer/employee," to evoke professional standards and interpretations. That sounds obvious, but it can get complicated quickly because while data sharing agreements are between (at least) two parties they also typically limit the way that those two parties can share data with others. In addition, there are ways that data sharing agreements can, explicitly and implicitly, create data rights for groups who are not party to the agreement.

## 2. Preamble

Preambles are optional in most data sharing agreements, but they are especially common between public institutions and between parties whose right to make an agreement is based on another relationship. Preambles can cite the founding authority of the parties, the events that led up to the agreement, or the mutual priorities and intentions of the parties. Private data contracts often include a provision in which the parties explicitly commit to having the necessary authorities to make the agreement.

### a. Purpose + Scope

WHEREAS, the parties desire to **[DATA SHARING/DATA USE]** in order to **[PURPOSE OF DATA RELATIONSHIP]**.

 Purpose. The purpose of an agreement is important because it is often the frame for, or a condition of, data sharing authorities, meaning that data shared between the parties is only valid if it's used for the purpose of the agreement. One of the main distinctions in understanding data management is the difference between 'purpose' and 'use'. The purpose of a data sharing relationship is, essentially, an intention or desired outcome - and so the validity and quality of data management and sharing can be judged based whether it contributed to the desired outcome. By contrast, a 'use' of data is more like an activity or action, which may or may not fall within a purpose - but the reason to include a purpose is to make it easier for courts to understand whether a specific action was intended by the parties.

As mentioned above, some data sharing agreements manage this by articulating the intended purpose with broad language, like 'the management of marine species in federal waters', whereas others focus on comparatively specific goals, like addressing ' bycatch of protected living marine resources.'

 Scope + Data Description. There are a range of ways to frame the scope of a data sharing agreement, each of which has an implication on impact and sustainability. Scope also, as importantly, often has a determinative impact on what data can be shared, under which circumstances, and by who.

WHEREAS, this Data Sharing Agreement (Agreement) is entered into by and between **[PARTY ONE]** and **[PARTY TWO]** to establish the **[CONDITIONS NECESSARY]** of **[TYPE OF DATA]** for the purpose of **[INSERT DESCRIPTION OF PURPOSE]**. This Agreement is only for sharing of the data described below and is a no cost agreement between the parties.

### b. Statement of Authorities (most common in public institution agreements)

WHEREAS, **[BASIS OF PARTY ONE LEGAL LICENSE/AUTHORITY TO POSSESS AND SHARE DATA FOR THIS PURPOSE]** grants **[PARTY ONE]** the license to participate in, share, and otherwise use **[DATA]**;

WHEREAS, **[BASIS OF PARTY TWO LEGAL LICENSE/AUTHORITY TO POSSESS AND SHARE DATA FOR THIS PURPOSE]** grants **[PARTY ONE]** the license to participate in, share, and otherwise use **[DATA]**;

 Conditional Rights (i.e - Confidentiality): Conditional rights are limitations on data sharing that are contingent on context. For example, confidentiality clauses typically create strict prohibitions on sharing data as long as it isn't common knowledge, but if the protected information becomes public, then those prohibitions disappear. Parties entering into data sharing agreements often possess data based on contingent rights, which is one reason to be clear about each party's authorities and their limits.

**[Replicate provision for each implicated/applicable authority necessary to support the scope of the data sharing relationship and its attendant limitations/conditions]**

NOW, THEREFORE, **[PARTY ONE]** and **[PARTY TWO]** hereby agree that their interactions shall be guided by provisions of this Agreement.

## 3. Term/Period of Agreement

The period of this Agreement shall be in effect from the date of last signature for a period of **[INSERT TIME PERIOD]**. This Agreement may be extended pursuant to **[AMENDMENT]** by consent and signature of both parties and may be terminated in writing by either party pursuant to **[TERMINATION]**.

 The term of an agreement is how long the parties expect to be involved in the data sharing relationship. The key distinction to observe in negotiating the term of data sharing agreements is the term of the agreement, as opposed to the term of the rights conferred through the agreement. Many data sharing agreements don't focus on a single data point or database, but rather create ongoing data transfer as long as the agreement is in place. So, the term of the agreement itself bounds the amount of time that the parties share access to data, as opposed to limiting how long the parties can use the data shared under the agreement. This is in contrast to data-specific licenses and limitations, which determine how long, and under what conditions, the parties can share data. The term of the agreement is the time during which each party collects and shares data.

## 4. Data

### a. Definition

Data sharing agreements are shaped by both the place and time they are created, as well as whether they focus on the data being shared, the activity being pursued, and/or the intended impact or outcome of the data sharing relationship. For example, in the United States, data is often regulated by whether or not it is personally identifying, and therefore requires a formal agreement to share, whereas many assume that, if data is not personally identifying, then it may not require a formal agreement to be shared. That assumption creates challenges if parties want to share data internationally or to combine the data with other data that would then make it personally identifying. Here, the point is that an agreement that references "personally identifying data" bases the scope of the agreement on the characteristics of the data.

By contrast, data sharing agreements can frame the scope around provenance (source), or supply chain of data collection. In this case, as long as the data is collected and transferred legally, then the parties are free to use the data however they wish. While this type of frame ensures continuity of rights in important ways, it can also be abused based on the contextual nature of data's value and risks. Another common approach to defining the scope of data sharing agreements is based on the common purpose, mandate, and/or role of the agreeing parties - usually with language like "parties agree to share all available data in order to rebuild overfished stocks of highly migratory species in the Pacific." In this approach, the volume and type of data shared is determined almost exclusively by its perceived utility to achieving a particular goal.

These approaches are not mutually exclusive and, in fact, can be used as layers so that parties, for example, are only able to share data collected in particular ways if it is useful to a shared purpose. Each approach to defining the frame of the data collected highlights the priorities, scope, and clarity of the parties involved and the use they envision.

**[PARTIES] agree to use information for [INSERT DESCRIPTION OF PURPOSE/GOAL OF DATA]. For the purposes of this agreement, data refers to [DEFINITION OF DATA].**

The most important value to seek in data definitions in sharing agreements is clarity. Because data is cheap and easy to copy, share, transfer, and make derivative products out of most agreements grant expansive rights to data. It is common to see terms like "global," "perpetual," and "royalty-free," in data agreements, allowing parties to use shared data, however they please. That unlimited use makes it very difficult to create contextually relevant rights protections, which increases data use but decreases the amount of control and context data owners can exert about how their information is used, even if it's against them.

Non-Disclosure + Confidentiality. In most relationships, the parties exchange private information: plans, communication about and observation of events, or commercial secrets. Provisions can be drafted to apply to data shared during a relationship, which typically enables the parties to share data internally as necessary to fulfill the main objective, like eligibility for a fishing license, with the expectation that the data is not used outside the scope of that relationship, and is deleted or returned at the end of the relationship.

Intellectual Property. In relationships where data is intellectual property (IP), the focus is typically on three core areas: (1) the ability of each party to use and create value from the data, and how, if at all, that value should be distributed; (2) the rights of the parties to prevent each other from realizing the value (and, in a much smaller number of cases, avoiding the risk) of data use and sharing; and (3) the act of publication, sharing, or granting additional license to the data as the core model of use. There are too many available models of IP agreements to try and draw best practices beyond the benefits of clearly articulated, bounded data uses and rights - as opposed to open-ended licenses which, initially, enable actors to exert their interests unfettered, but rarely help stakeholders or parties resolve disputes.

Applicable Law (Reporting). Beyond fisheries, specifically, there are a number of other adjacent regulations with an effect on data, such as business reporting requirements or data retention. It's important to understand those laws and how they may constrain the parties' ability to make guarantees, especially about independence from law enforcement.

**Description of [PROTECTED/CONFIDENTIAL] Data:**

**[INSERT TECHNICAL DESCRIPTION OF DATA]**

If the Data are used for any other purpose other than authorized under this Agreement, **[INSERT DESCRIPTION OF PENALTY FOR VIOLATION OF STANDARD/PROTECTION/CONFIDENTIALITY]**, and this agreement will be terminated immediately.

In addition to agreeing what the parties should do, data sharing relationship negotiations often consider how serious the consequences should be if they don't do the things agreed. While this is typically the role of a dispute resolution system, parties can indicate the general size and type of harms that might arise from breach - which can help courts understand the original intent of the parties. This also means that penalties are typically limited to the kinds of things that courts can compel, which are: (a) financial - imposing a cost or fine for a particular action; (b) injunctive - a much more rare, but rule-based prevention on a specific action; and (c) freedom - typically limited to criminal matters, courts can jail and, in very limited circumstances, restrict specific freedoms.

### b. Data [Operational Acquisition Conditions/Means of Collection]

All **[DATA]** collected and transferred by **[PARTIES]** shall be done according to **[OPERATIONAL CONDITIONS + TECHNOLOGY STANDARDS]**.

There are a range of competing legal and technical standards for defining the "best," approach to data management. Standards are rules of practice that are often independently certified trade organizations or public institutions. In the United States, the National Institute of Science and Technology, for example, maintains standards for data integrity, security, and sharing. In addition, there are large, international technical standards federations, like the International Organization of Standardization (ISO) and the Institute of Electrical and Electronics Engineers (IEEE), both of which generate standards for how to manage and share data. Both bodies are known for their rigor so setting standards is less about picking the 'right' standards, than about how to effectively monitor and enforce those standards. Certification bodies are one way, but they do more to ensure general practice than prevent specific harms.

Beyond ensuring data management professionalism, data sharing agreements typically include assurances as to the quality or fitness of a data or technology product for the articulated purpose. Warranties are the guarantees that a party makes about the quality and necessary conditions for use of its products. By articulating a guarantee, the party assumes legal liability in case the product does not meet those standards and, by implication, disclaims liability for things that fall outside of those uses. While traditional commercial contracting often features warranties as a sign of quality, digital and data contracting often make no guarantees about data quality or, even, legality of collection. *One of the reasons for this is that data can be used for many things, and so warranties that limit future use are seen as anathema to the technology industry's focus on unbounded innovation. Warranties, indemnity clauses, and disclaimers are all important ways that data sharing agreements frame and, more commonly, limit the liabilities that parties assume for the integrity of the data sharing ecosystem.*

In data sharing agreements, standards and warranties are a way to set expectations, and to articulate the consequences of failing to meet them. When misused, standards can be a way to generalize away liability for failing to prevent predictable harms - like poor due diligence or oversight. Data standards are a way to avoid being deterministic about how each party manages their work, while upholding quality.

### c. Constraints on Use of Data

The **[PARTIES]** agree that each party is **[ARTICULATION OF OWNERSHIP/ONGOING AUTHORITY]** of **[DATA].** The **[PARTIES]** will in all respects treat **[DATA]** in accordance with all procedures reasonably necessary to protect **[PARTIES]** rights therein, consistent with the legal authorities available **[Optional: DIRECT REFERENCE TO AUTHORITIES]**. In the event that either party receives a legally valid request that challenges or requires exception to these rights, the parties agree to give the other party as much notice as possible.

1. **[PARTY/PARTIES]** shall only use the **[DATA]** for the purposes identified in **[PURPOSE + SCOPE]**.

Data is typically shared not only through agreements but also through licenses. Licenses are the specific conditions, uses, and rights under which data is shared between parties and the way they frame data sharing rights and limits can give parties the right to share data with others.

2. To the extent permitted by applicable laws and regulations, **[PARTY/PARTIES]** shall restrict disclosure of the **[DATA]** solely to those who have a need to know such information in order to accomplish the identified purposes. **[Optional: Specific list of additional users/individuals, organizational types, etc.]**

Non-Disclosure + Confidentiality In most relationships, the parties exchange private information -plans, communication about and observation of events, or commercial secrets. Provisions can be drafted apply to data shared during a relationship, which typically enables the parties to share data internally as necessary to fulfill the main objective, like eligibility for a fishing license, with the expectation that the data is not used outside the scope of that relationship, and is deleted or returned at the end of the relationship.

3. **[PARTY/PARTIES]** shall advise such individuals, before they receive access to **[DATA]**, of obligations under this Agreement. **[PARTY/PARTIES]** shall take steps to ensure that all such individuals shall comply with the provisions of this Agreement.

Assignment: Assignment clauses are what enables parties to convey their interest or rights to other parties - either directly, through subcontracting, or indirectly, based on a change in circumstances, like bankruptcy or death. The benefit of flexible assignment clauses is that they enable adaptable data workflows; the downside is whether and how the conditions of the underlying agreement will apply to assigned parties.

4. **[ADDITIONAL CONDITIONS OF USE AS NECESSARY]**
5. This Agreement imposes no obligation upon **[PARTY/PARTIES]** with respect to information that is received outside the scope of this agreement and/or is required by law or a court to be disclosed.

One of the main reasons that data sharing agreements are so important is that it is extremely hard to build perfect security or unhackable infrastructure. Rather than expect perfection, an increasing number of data sharing agreements rely on parties to take 'reasonable' measures to collect, manage, use, transfer, store, and delete data, among other things.[1] While these activities are supported by a growing range of technical, legal, and professional standards - they still create the unavoidable potential for risk.

One of the main functions of data sharing agreements is to clarify and apportion the risks, or potential for risks, posed by data transfers. While there are a predictably wide range of practices and their nuances are heavily contextual, most data sharing agreements vary based on the degree of liability the parties take, how mutually or equally they share the blame for breach, and what kinds of assurances or steps the parties are able to make to secure their systems.[1]

In most professional and commercial contracting, standards of professional conduct are an important way to ensure that each party takes reasonable care in the tasks they perform. One of the trends in data sharing agreements, and technology contracting more generally, has been a move away from most forms of liability, including professional standards of conduct, as well as basic guarantees about how the data or project works. Whereas many industries make specific guarantees or warranties about their products, often with the caveat of the conditions and maintenance required for their use, many commercial data sharing agreements broadly disclaim the quality and accuracy of the data, in addition to broadly indemnifying each other. The removal of those traditional liabilities can also reduce incentives to build or articulate professional standards for data management.

**d. Professional Standards and/or Guarantees [Security, Protection, Privacy, etc.]**

1. To protect the **[DATA]** from unauthorized physical and electronic access and to ensure the confidentiality, availability and integrity of all data shared, **[PARTY/PARTIES]** shall meet or exceed industry best practices for physical security, data security, network security, and access controls, both technically and procedurally, including compliance with **[PROFESSIONAL/TECHNOLOGY STANDARDS.]**

2. **[STANDARDS OF SYSTEM MAINTENANCE]** The **[PARTY/PARTIES]** agree to ensure that all systems used to manage, or otherwise interact with **[DATA]** under this agreement are maintained **[REGULARLY/SPECIFIC PROFESSIONAL STANDARD]**

3. **[COMMUNICATION/NOTICE IN THE EVENT OF BREACH]** In the event that **[PARTY/PARTIES]** become aware of a breach of any security, confidentiality, or other management commitment made under this agreement, they will notify the other party within **[PERIOD]**.

Most data sharing agreements also enable 'termination for cause' - which gives a party the right to unilaterally end the agreement and/or the rights conferred through it, based on a breach of the agreement. Termination for cause provisions vary based on how much notification the parties need to give and whether there are opportunities to fix the problem in good-faith.

4. **[COOPERATION/INDEMNITY FOR SYSTEM INTEGRITY]** The **[PARTY/PARTIES]** agree to share data, service, and other resources, as possible, toward supporting the investigation, pursuit, and restitution for any damages resulting from breaches to any data management commitment made herein.

5. **[PARTY/PARTIES]** retain the right to assert a claim for any damage caused or costs incurred because of the failure of **[PARTY/PARTIES]** to perform an obligation under this Agreement or the negligent acts of its personnel, employees, or contractors which results in an unauthorized disclosure, release, access, review, or loss, theft, misuse or destruction of confidential data. The treatment of such a claim by **[PARTY/PARTIES]** shall be reviewed and adjudicated pursuant to **[APPLICABLE LAW]** and regulations; provided, however, that nothing herein shall be construed to be an undertaking to indemnify or hold-harmless any person whether or not a party to this agreement.

Beyond ensuring data management professionalism, data sharing agreements typically include assurances as to the quality or fitness of a data or technology product for the articulated purpose. Warranties are the guarantees that a party makes about the quality and necessary conditions for use of its products. By articulating a guarantee, the party assumes legal liability in case the product does not meet those standards and, by implication, disclaims liability for things that fall outside of those uses. While traditional commercial contracting often features warranties as a sign of quality, digital and data contracting often make no guarantees about data quality or, even, legality of collection. One of the reasons for this is that data can be used for many things, and so warranties that limit future use are seen as anathema to the technology industry's focus on unbounded innovation. Warranties, indemnity clauses, and disclaimers are all important ways that data sharing agreements frame and, more commonly, limit the liabilities that parties assume for the integrity of the data sharing ecosystem.

In data sharing agreements, standards and warranties are a way to set expectations, and to articulate the consequences of failing to meet them. When misused, standards can be a way to generalize away liability for failing to prevent predictable harms - like poor due diligence or oversight. Data standards are a way to avoid being deterministic about how each party manages their work, while upholding quality.

**e. Disposition of Confidential Data**

All **[DATA]** received by the **[PARTY/PARTIES]** under this Agreement that is no longer needed by **[PARTY/PARTIES]** for the purposes described in **[PURPOSES SECTION]** shall be handled in the same manner as noted below **[TERMINATION]**.

## 5. Party Authorities

### a. Amendments and Alterations to this Agreement

This Agreement constitutes the entire agreement between **[PARTY/PARTIES]** for the purpose of using **[DATA]** for the **[PURPOSE]**, and no other statements or representations, written or oral, shall be deemed to exist or to bind the parties. **[PARTY/PARTIES]** may amend this Agreement at any time, including extending the duration of the Agreement, by mutual written agreement **[TERMS OF AMENDMENT]**.

This is typically managed through the 'Amendment' provision in data sharing agreements, and it sets out the conditions for changing the rules. The ability to change the rules is as important as the initial agreement, as whoever can change the rules can subsequently alter its terms. The primary difference between Amendment provisions comes down to how mutual the authorities are, and how granular the changes. For example, a number of technology platform contracts give the technology provider the unilateral authority to change the terms of service, whereas more directly negotiated data sharing agreements often require mutual, written agreement to amend the terms. One way to manage the potential abuse of those authorities is to limit the scope of the Amendment provision to minor technical changes, or to ensure that parties do so 'in good faith' and with ample notice to the other party. Ultimately, Amendment authority is one of the most important elements of bridging the gap between a single instance of data sharing as an activity, and the broader impacts it has on, and ways it reflects, the underlying relationship between the parties.

### b. Termination of the Agreement

**[PARTY/PARTIES]** may terminate this Agreement at any time, for cause. **[PARTY/PARTIES]** may terminate this agreement for any reason with a 30-day written notice. The Agreement will terminate at the end of the time period identified in **[TERM]**, unless extended pursuant to section **[AMENDMENT]**.

When termination happens in the normal course of a data sharing relationship, it's either because of the expiration of the Term, or the parties have fulfilled the requirements of the termination provision. The primary difference between termination provisions, like amendment authorities, are how equally or mutually agreed they're required to be. Most data sharing agreements also enable 'termination for cause' - which gives a party the right to unilaterally end the agreement and/or the rights conferred through it, based on a breach of the agreement. Termination for cause provisions vary based on how much notification the parties need to give and whether there are opportunities to fix the problem in good-faith.

In the event this Agreement is terminated, or **[PARTY/PARTIES]** cease operation, **[PARTY/PARTIES]** shall remove general user access to the **[DATA]** and related materials. **[PARTY/PARTIES]** disposition of such records shall be in accordance with applicable law. **[PARTY/PARTIES]** will destroy or return all **[DATA]** covered by this Agreement as well as any backups and all related materials that contain any portion of the **[DATA]** to the originating party.

**c. Assignment of the Agreement**

**[PARTY/PARTIES]** may outsource, assign, subcontract, or otherwise transfer the rights granted under this Agreement at any time, with the written, mutual agreement of both parties.

Assignment: Assignment clauses are what enables parties to convey their interest or rights to other parties - either directly, through subcontracting, or indirectly, based on a change in circumstances, like bankruptcy or death. The benefit of flexible assignment clauses is that they enable adaptable data workflows; the downside is whether and how the conditions of the underlying agreement will apply to assigned parties.

## 6. Choice of Law/Operational Treatment

Data sharing agreements typically point to the institutional and external authorities the parties will defer to resolve disputes. This usually happens on two levels: the choice of law clause and, often, a dispute resolution clause (sometimes referred to as an arbitration agreement). The choice of law clause affirmatively picks a legal jurisdiction for the agreement, which means that jurisdiction's laws govern the relationships and, if the parties disagree, they will go to court in that jurisdiction. Choice of law clauses are important because they import the applicable laws of that jurisdiction and, essentially, subject the parties to the influence of its public institutions.

**a. Compliance with Applicable Laws and Regulations**

**[PARTY/PARTIES]** shall comply with all applicable state and federal laws and regulations protecting the **[DATA]** pursuant to this Agreement. The validity, interpretation, construction and performance of this Agreement shall be governed by the **[CHOICE OF STATE/ADJUDICATING JURISDICTION] [OPTIONAL: more common for federal/public institution agreements:** "but, giving effect to federal laws governing the behavior of purpose/fisheries data management."**]**

Applicable Law (Reporting) Beyond fisheries, specifically, there are a number of other adjacent regulations with an effect on data, such as business reporting requirements or data retention. It's important to understand those laws, and how they may constrain the parties' ability to make guarantees, especially about independence from law enforcement.

### b. Dispute Resolution

Except as otherwise provided in this Agreement, disputes arising between **[PARTY/PARTIES]**, shall be resolved through discussion and negotiation **[Optional: Arbitration Agreement]**. If the Parties are unable to resolve the dispute through good faith discussions, the Parties reserve the right to seek legal remedies as appropriate.

Dispute resolution clauses are what governs everything leading up to the decision to go to court. There are a range of dispute resolution approaches, ranging from the parties giving each other good-faith notice and an opportunity to respond to forcing parties into binding arbitration. Dispute resolution clauses, when well-designed, clearly articulate a process for identifying and managing disagreements collaboratively and an escalation pathway that preserves as much of the value of the relationship as possible. When parties decide that they are unable to resolve a dispute themselves, they have the choice of whether to go directly to court, or to work with alternative dispute resolution services - like mediation and arbitration. In these cases, parties typically negotiate and enter into a second contract or agree to an arbitrator's ruling on how they will resolve the dispute. While there are a range of contexts and decisions that go into resolving disputes coming from data sharing agreements, it rarely serves the best interest of either party to forego their right to use the courts.

### c. Notices

Any notices made in connection with this Agreement shall be in writing and **[MEANS OF COMMUNICATION]** to the following individuals and addresses:

The means of communication can be anything that the two parties mutually agree upon, such as: formal letter, email, fax, phone call, text message or some combination.

For **[PARTY ONE]**:

**[INSERT POINT OF CONTACT]**

For **[PARTY TWO]**:

**[ INSERT POINT OF CONTACT]**

### 1.0 [OPTIONAL: Independent Capacity]

The employees or agents of each party who are engaged in the performance of this Agreement shall continue to be employees or agents of that party and shall not be considered for any purpose to be employees or agents of the other party. By the signatures of their duly authorized representative below, the **[PARTY/PARTIES]** agree to all of the provisions of this Data Sharing Agreement.

Independent Parties clauses protect the parties' individual autonomy, toward ensuring that the Agreement is not construed to give them the right to make representations or commitments on the others' behalf.

**NOAA FISHERIES**

BY: _____ BY: _____

TITLE: TITLE:

DATE: _____ DATE: _____

# Appendix B: Data Sharing Language Repository

**Preamble:**
This repository has been designed as a living document in support of the other two parts of this package: the resources section and the annotated data sharing agreement. This is also a starting point intended to grow and evolve, both in terms of the definitions for certain terms and the addition of new terms, as well as the removal of any terms not deemed useful. Two items of note:

## 1. Definitions and Terms

The definitions and explanations in this document are not from a dictionary nor are they the only possible interpretations of the words. Given that data is so reliant on the context it comes from and is used in, we have contextualized the terms to be practically useful in the HMS PSG fisheries context and, where possible, to NOAA. Sharing this documentation with the parties involved in a data sharing discussion may help improve the process by creating shared vocabulary or at least language to discuss areas where clarity of purpose is vital. These conditions inform a more defensible negotiations process.

## 2. Organizing Structure

The structure of this document organizes the terms in four groupings that relate to different conversations that occur in the course of the creation of a data sharing agreement. The groupings are as follows, and a list of the terms is included in the table of contents of this document so that you can see at a glance if there is an entry for the topic you'd like to read more about.

**Data Sharing Language Repository Categories**
Parties and Authority
Data Terms in Focus
IT Systems & Technology Products
Miscellaneous

## Agreement

An agreement is a recognized or legally binding set of terms, often in the form of a contract. Agreements have various amounts of formality. The amount of formality and congruence with a regulatory standard of court determines how enforceable it is.

A 2019 report from the Duke Nicholas Institute and the Internet of Water, "Assessment of Federal and State Agreements with Data Organizations," (the "Duke Report") is an assessment of federal and state Agreements with data organizations[24]. This report is a helpful guide to the various types of data sharing agreements that are in use between federal and state agencies and data organizations.

As it relates to data sharing for fisheries management, agreements currently may take the form of an MOU or a service-based data license agreement. Another type of data sharing agreement is a consent-based license, which is a mutually constructed agreement that enables two [or more] parties to design a bounded and specific contract through direct data licensing for a certain set of purposes. Over time, a number of different agreements have been and are used to define and direct data sharing across NOAA, and these are not always referred to as "agreements." For example NOAA Administrative Order 216-100 "Protection of Confidential Fisheries Statistics" refers to "all binding forms of mutual commitment under a stated set of conditions to achieve a specific objective."

## Authority

An authority is a right that is enforced by a public institution. A "right", such as rights of people who share their data, or the rights of the public to access data, are the recognition by a system of authority. When actors make data sharing agreements, the agreements point to the authorities that we expect to enforce them, as well as the authority for the rights to underpin them. In the case of Magnuson-Stevens, the federal government's authority is used to define if and how data must be shared.

## Direct Authority + Consent

In a data sharing agreement for a fisheries case, an independent fisherman represents a party with direct authority. They are able to engage in a contract or data sharing agreement directly with another party based on consent. A representative of a fishing organization may not have direct authority and may need approval from senior staff and/or their Board to commit the organization to a contract. In addition, most contract law is based on the idea of freely negotiated, informed consent. While those terms are misused a lot, consent-based authority assumes that the parties' involved proactively choose to participate, as opposed to being required to do something by law, or as a condition of a more important aim.

[24] "Assessment of Federal and State Agreements with Data Organizations", Duke Nicholas Institute for Environmental Policy Solutions, Internet of Water. November 2019.

## Regulatory + Constructive Authority

A regulatory authority is an autonomous authority or agency established by a federal, state or provincial government. The regulatory authority for fisheries management is the U.S. Department of Commerce which houses the National Oceanic and Atmospheric Administration (NOAA). The court that enforces the terms of a data sharing agreement is determined by the choice of law clause, and the jurisdiction chosen by the parties entering into the agreement.

## Consent + Contingent Authority + Consent-based License

Informed consent is the foundation of voluntary data sharing. Informed consent can involve negotiating or setting expectations about the actions of all parties: what will you allow others to do with data that you share with them, and what will you be allowed to do with data that is shared with you? And, in addition, what previously granted consents – for example, those given by fishermen or other data subjects – frame or bound the authorities of the parties to the data sharing agreement. Whose interests, essentially, convey with the data?

This is a case where your role and the ways that you use data can help you create a very clear point on what you seek consent for. What kind of a report do you need the data for? What is the use? It is vital to push past the “who” you allow data sharing with and dig into the “what” is allowed to be done with the data - how can it be used? How can it not be used? Stating these uses explicitly can be part of defining clear consents.

If you take the idea of what you're seeking consent to do and make it very clear, you will create a well-signed path for you to follow if and when the use of data changes. Say the format of your quarterly report changes to include a new economic analysis - are you still using data in the same way? If not, perhaps it is time to revisit the agreement to check-in?

Consent can be expressly granted or implied. “Express” consent means that consent is given directly, either as spelled out in the agreement, or in response to future requests made by voice or in writing. “Implied” consent is manifested by signs, actions, or facts, or by inaction or silence, which raise a presumption that the consent has been given. Clarity in data sharing agreements is well served by the creation of express consent.

Consent as recognized by the law is informed consent. The law struggles where there are open-ended and therefore unknowable consents between two parties. An example of open-ended consent is when a contract or law says who is allowed to access data but stops short of explaining what it can be used for. If something is open-ended then it is considered by the law to be less informed. This is why clearly describing the data use you are consenting to, and for what purpose, is part of moving towards a more accessible and easy to use data sharing approach.

One of the key concepts of consent and contracting is that they are ‘freely’ entered into. Purely volitional, consent-based licenses are the simplest version of the authority necessary to create data sharing agreements, and can grant exceptionally flexible rights and uses. By contrast, there are a number of data sharing relationships that are compelled, or required components of larger relationships - for example, in regulatory reporting. A fisherman agrees to data sharing as a condition of their publicly granted license to fish, so the consent to data sharing is contingent on the things strictly required to provide that license. Contingent consent is valid and a normal condition of data sharing relationships not primarily defined by data, though these typically lack the flexibility of consent-based licenses. Placing the idea of consent (see above definition) into an

agreement is the pillar of consent-based data licensing. One of the values of basing data sharing on consent is that the law gives people a lot of flexibility in what they can agree to.

In data management and sharing practice, there's a tension between how limited and specific the authorities granted in a license are, and how 'informed' one can be about the potential uses of data shared. These debates are most common with open-ended licenses and unilateral authorities in data sharing agreements, which can make it difficult to bound how that data will be used.

## Amendment Authority

One piece of a successful data sharing agreement is a clear and well-defined path to change. So many factors can change over time - technology, politics, policy, the environment - that it's vital to define how changes would be made to a data sharing agreement, and who holds the power to do so. This is called an "amendment authority," and it refers to the agency or persons that have the authority to amend an [agreement]. In consent-based licensing, the parties to the agreement can define who has amendment authority; in some cases deferring to one party, but more commonly requiring some form of mutual agreement.

## Arbitration Agreement

An arbitration agreement is typically a clause in a contract that requires the parties to resolve their disputes through an arbitration process. It is a form of alternative dispute resolution.

## Authorized Use(s) + User(s)

The idea of 'authorized' use and users means a use/users that are allowed by whatever authority compels the agreement. Authorization of data management and sharing is usually based on a combination of factors, including that the activity (use) and/or the person or institution managing the data (user), and most often a combination of both.

More formally, authorized uses are: [From NAO 216-100] "that specific use authorized under the governing statute, regulation, order, contract or agreement," and authorized users are "any person who, having the need to collect or use confidential data in the performance of an official activity, has signed a statement of nondisclosure affirming the user's understanding of NMFS obligations with respect to confidential data and the penalties for unauthorized use and disclosure."

Authorization requires both valid rights and systems of authority, as well as the ability to prevent unauthorized use. One challenge is that effective authorization can also require the monitoring and enforcement capacity to ensure that the use is limited to that which is authorized.

## Breach

A breach is when either party fails to observe and uphold the terms of an agreement. It is defined by local law. The associated penalty for any breach is dependent to a certain extent on the limits of the jurisdiction in which the parties or a party bring a case. Essentially, the terms related to a breach represent how parties agree to hold each other accountable with how they construct the agreement.

## Contract

A contract is a legally recognized agreement to do or not to do a particular thing.

## Disclaimer

In the context of a data sharing agreement, a disclaimer may be used to avoid any liability about the quality or methodology of data. When someone “disclaims” something, they are absolving themselves of a right or a claim. For instance, in a data-sharing agreement, one could seek to disclaim all liability for quality and completeness of the data that is being shared and as such, the data may have dangerous flaws embedded within.

By contrast a “warranty” does the opposite, and clearly states what purpose the data should be used for while committing the parties to being liable for making the agreement work.

## Enforcement

In the context of a data sharing agreement, enforcement emanates from the authority that defines the terms of the agreement, and then by the authority that enforces it. All data sharing agreements should grapple not only with what happens if the parties breach the agreement, but also how the aggrieved party will be able to enforce the terms of the agreement if the other party refuses to cooperate. Doing something that goes against the terms of the agreement may be the trigger for an enforcement action. Private data sharing agreements are typically enforced by civil authorities. Under the auspices of regulatory authority, public institutions can also enforce breaches of data sharing agreements - whether through their own leverage, such as controlling licensing, or via external authorities, like courts. Importantly, contracts do not get enforced unless they are broken, and parties are not always aware of their breach, nor do they always have the resources to seek redress in court. That is, ultimately, the argument for arbitration agreements and alternative mechanisms.

## Indemnity

Data sharing agreements often grant broad indemnity - essentially, relieving the other party of legal liability. The most common argument for indemnity in data sharing relationships is that data is so cheap and easy to share broadly that it should be as unencumbered with liabilities as possible. In a data sharing agreement, indemnity can be sought to pre-empt recourse for the use of data. The most common indemnity sought is to hold the other party harmless for a certain action, should it happen. Much like a disclaimer, indemnity is a proactive, explicit release of liability.

## Jurisdiction

In a data sharing agreement, “jurisdiction” refers to the legal authority or power to decide an issue. Another way to think of the term is the answer to the question “who decides?” when assessing if the terms of an agreement are being followed. The term may also refer to the regulatory geography of an agreement, noting which state or (other) jurisdiction the terms of the agreement are bound by. Because rules vary state to state, Jurisdiction is an important matter to have clarity on.

## License

In a data sharing context, a license is a tool used to describe the permissions and relationship of a data owner and a data user. The terms of the license can be defined through a data sharing agreement or contract. A license is a way to capture the things that a data user is allowed to do. If you consider an open-ended data license, it may include terms that say data can be used by a party for the general activity of fisheries management which would permit a wide range of uses. In contrast to an open-ended data license, a consent-based license might indicate that data can be used for the generation of three specific analyses in a certain geographic region, for governmental purposes, and several other research processes so long they are not commercial purposes. In both cases, the license is the tool used to capture the idea of a permission of use between two parties.

## Memorandum of Understanding ("MOU")

Memorandums of Understanding (MOUs) are more informal collaboration agreements than contracts and agreements. In the context of data sharing, MOUs have become a commonplace way to move from fully informal and unwritten agreements towards something more formal.

In a data sharing context, an MOU usually describes, in broad and general terms, a data sharing relationship managed cooperatively by two or more agencies or entities. MOUs generally do not include specific information regarding data use or specific boundaries for use. MOUs tend to have less specificity than contracts. MOUs can help stakeholders agree to a high-level set of standards and norms and may signal intent to develop a more formal data sharing relationship. This type of agreement is common at NOAA.

## Non-Disclosure Agreement ("NDA")

In the context of data sharing, a non-disclosure agreement signals that data cannot be disclosed to anyone beyond the parties and terms describing the agreement. One example is the Non-Disclosure Agreement for Confidential Vessel Monitoring System Information used by the NMFS Office for Law Enforcement. This means that NMFS is not allowed to share confidential data with an individual or entity unless they are a grantee or direct service provider.

## OPEN – "Open, Public, Electronic, and Necessary" Data Act

The Open, Public, Electronic and Necessary (OPEN) Government Data Act (2017) provides a government-wide mandate for federal agencies to publish all of their fisheries information as open data using standardized, non-proprietary formats.

## Open Data Policy

The US "Open Data Policy" refers to an Executive Order, "Making Open and Machine Readable the New Default for Government Information" from 2013. The policy was created to expand public access to the results of federally funded research. You can learn more about the Open Data Policy and how it relates to a Strategy for American Innovation on the Digital.gov website. In February of 2015, NOAA responded to this direction with the release of their plan for increasing public access to research results.

## Open-Ended License

An open-ended license, in the context of a data sharing agreement, does not specify details about the use of data. It provides general guidance for the use of data but leaves a lot of room for interpretation of what might constitute permissible use.

## Parties

In terms of a data sharing agreement, the parties are the organizations (or individuals) creating a formal relationship that defines how and why they each have access to data and the terms of its use. Relevant information to be considered when negotiating data sharing terms relates to the authority, or power, of each party and how their formal engagement in the agreement translates into operational use and accountability. For example, in the case of NMFS, when an agreement is entered into by NMFS, how do staff in different positions learn about and uphold the terms that the organization has signed up for? It is important to map how being party to an agreement plays out in practical and operational terms in order for data sharing to be a literacy shared and enjoyed by a range of actors across an organization. It is also important to know how to uphold the roles and responsibilities defined by the terms of an agreement. Some examples of parties in the fisheries context are: fishermen, a vertically-integrated tuna company, NMFS regional offices, academic researchers, or any other stakeholders with an interest in fisheries management.

## Right

A right is a legally recognized and enforced interest. In the creation of data sharing agreements and contracts, rights are written down and assigned to the parties to the agreement. Knowing what you have rights to do (and not to do) with data that you access can go a long way to creating clarity and confidence in data sharing. In terms of fisheries management, the Magnuson-Stevens Act establishes the public interest in fisheries conservation and support as a condition of granting the private right to fish. If the law does not specify a right, it can be harder to negotiate. In order to develop a robust and satisfying data sharing agreement, all parties need to agree which rights you are negotiating *from* - what is the baseline and the starting point. What rights do each party have in the service of their respective livelihoods, whether the oversight of a public good or a commercial operation? This ‘existing rights’ foundation sets the stage for what protections and relationships are negotiated in an agreement.

## Service-Based Data License

A “service-based” data license specifies that the user of data only receives access to data in order to provide a particular service. This could be a data license agreement type used with a vendor in order to leverage their analytical software. It means that the service relies on the data and it also constrains the use of the data to that particular context - meaning that you only get the data license if you are part of some service.

## Term/Period of Agreement

In a data sharing agreement, the “Term” or the “Period of Agreement” refers to the time period for which a data sharing agreement is valid. Time is an important feature of data sharing because as circumstances and technologies change, the conditions of data sharing may need to be examined. By specifying a term/period of agreement, it’s possible to build a review trigger into any contract, which will force the parties to the table to consider whether all is good with the

agreement, at which point it can be renewed, or if new context or situations have surfaced that would make a reassessment of the data sharing terms a helpful step. Building term agreements can also assuage fears of things that could go wrong by creating a period of testing or discovery which lowers the stakes on the decision to try a new data sharing agreement, or a new condition of use.

## Warranty

In the context of data sharing, the idea of warranty can be a helpful one to think about when moving from open-ended sharing of data to more precise and intentional use. When negotiating the sharing of data in an agreement, it could be considered the notes you keep in your head or the margins of your spreadsheets: when you share your data with someone, what is the context and additional information they need to know? And how can you help them contain the use of the data you are sharing with them to the right purpose? If you reverse this question, you can help get yourself the information you need about the data you receive - what are the questions you need to ask of your data source to use this data with confidence?

Generally, a "warranty" is a written statement that describes the conditions of use under which a product or service is guaranteed and intended for. "Warranties do two critical things. First, they require a manufacturer to specify how and where its product should be used and to run tests to ensure it performs under these conditions. And second, warranties attach liability to the risks created by a product to the user. Warranties are familiar enough to seem trivial, but they are an important part of how we structure the responsibility for the things we create, share and sell — including intangible and digital products."

### Data Terms in Focus

## Aggregate [or Summary Form]

Aggregate data is generally a data set that is a compilation of individual data records, without the individual records. For example, an aggregate data set may say how many fish were caught on a particular day without breaking it down as to where or by who, which are the individual data records that were used to create the aggregate data. Sometimes the term aggregate data is used interchangeably with anonymized data, which is incorrect, given that it is still possible to use other information to disaggregate and identify the individual data records used to create the aggregate data set.

NAO 216-100 - data structured so that the identity of the submitter cannot be determined either from the present release of the data or in combination with other releases.

## Anonymized Data

The intention of the term anonymized data is data or information that does not identify the people or sources it came from. While the term is used often as though it's possible to create anonymized data, it is not something that can be reliably accomplished in practice. There are technical processes that can be followed in an effort to achieve anonymity, some of which are referred to as "computational anonymity." And there are a range of mathematical models and approaches to

support this effort in practice and in design. Given the current state of both approaches and challenges with data anonymization, it is an important feature to address in data sharing agreements so that expectations are properly managed and the feasibility of anonymizing data is understood. It is important that people who undertake best efforts to follow processes or use computational means to achieve anonymization are held to the appropriate standard of care for these practices, and not beyond them.

In the context of data sharing, anonymizing data is best understood as a process that data is put through in an effort to mask its sources. This is done in a number of ways, and sometimes through computation. Both inside this industry and outside of it, many organizations and institutions are grappling with the idea of what anonymized data means. One of the major reasons to pursue anonymization is research, so that parties can share data sets with researchers and not worry about compromising the privacy of the sources (this relates to banks, health care providers, and a long list of examples).

## Confidentiality/Confidential Data

One growing tension with digital transformation is that the amount of things that are knowable by others outside of the purview of data sharing agreements has grown substantially. There is a great deal of tension around the term “confidentiality” that may have been recently exacerbated by the Open Data Act. Fisheries have long had public reporting requirements where data stewards agree to keep individual information secret while also publishing some aggregate or transformed data products openly. Whichever party offers the confidentiality, they must have some way to keep the data secure and private.

Private agreements (e.g. with regard to self-monitoring devices, technological tools to manage fleets and catches, and whatever is observable through other public data sets like satellite imagery) often have confidentiality requirements and boundaries as well. For instance, a fisherman’s ability to replicate the data that they give to NMFS or access from other services might diminish the potential of NMFSs promise of confidentiality under Magnuson as the boundaries of that confidentiality have thinned over time. Put another way, fishermen might be sharing data in their own right, which impacts NMFS’s capacity to manage the confidentiality of the data.

## Data

Given that data is the lifeblood of NOAA’s work managing a sustainable fishery system, being clear on what constitutes data is important. From NAO 216-100: “Data refers to information used as a basis for reasoning, discussion, or calculation that a person may submit, either voluntarily or as required by statute or regulation.”

In the NMFS Data and Information Management Policy Directive (renewed March 2013), the term "data" refers specifically to the alphanumeric values that are recorded as a result of a collection program or experiment, and to derived data that are the result of analysis or other synthesis. Data may be collected by NMFS staff, through contracts or grants, or in cooperation with the states and other partners. Data are not limited to recorded observations and measurements of the physical, chemical, biological, geological, or geophysical properties or conditions of the environment. Data may also be correlative data with related documentation or metadata, audio recordings, images, maps, photographs, or reports, and may include fishery-dependent and fishery-independent data, regulatory data, and other non-environmental types of data.

Specificity in terms of the data being shared and used in an agreement is important. A range of methods, from standards to metadata to descriptive methodologies, can be used to ensure that all parties to a data sharing agreement have an understanding of the data being considered in any data sharing relationship. It can also make sure that changes, or derivatives of the data, can be considered or discussed.

## Data Asset

From the NMFS Data and Information Policy Directive: Any entity that makes data usable. For example, a database is a "data asset" that includes data records. A data asset may be a system or application output file, database, document, or web page. A data asset also includes a service that may be provided to access data from an application. For example, a service that returns individual records from a database would be a data asset. Similarly, a website that returns data in response to specific queries would be a data asset. A human, system, or application may create a data asset. Data-sharing agreements are also a form of "data asset."

## Data Quality

The Data Management Body of Knowledge (DMBOK) defines data quality as "the planning, implementation, and control of activities that apply quality management techniques to data, in order to assure it is fit for consumption and meet the needs of data consumers." Since expectations about data quality are not always verbalized and known, an ongoing discussion is needed. Data quality depends on context and the data consumer's requirements.

## Data Steward

From the NMFS Data and Information Policy Directive: The individual who is responsible for establishing and maintaining the quality, integrity, documentation, and preservation of the data asset. A data-sharing agreement typically discloses who or what has the responsibilities of data stewardship.

## Enterprise Data

From the NMFS Data and Information Policy Directive: Data and information shared with an entity outside of the NOAA Fisheries Financial Management Center (FMC) of their origin. (For example, observer data shared between a Science Center and a Regional Office.) Enterprise data can include entities or issues that cross FMC boundaries (for example, data about vessels that fish in multiple regions). Enterprise data are also those data that are routinely aggregated to support decisions at a higher organizational level (such as catch and value data that is summarized in Fisheries of the United States).

## Environmental Data

DPSD - Are defined by NOAA Administrative Order (NAO) 212-15: Management of Environmental Data an Information as recorded and derived observations and measurements of the physical, chemical, biological, geological, and geophysical properties and conditions of the oceans, atmosphere, space environment, sun, and solid earth, as well as correlative data such as socio-economic data, related documentation, and metadata. Digital audio or video recordings

of environmental phenomena (such as animal sounds or undersea video) are included in this definition. Numerical model outputs are included in this definition, particularly if they are used to support the conclusion of a peer-reviewed publication. Data collected in a laboratory or other controlled environment, such as measurements of animals and chemical processes, are included in this definition.

## Information

From the NMFS Data and Information Policy Directive: Any communication or representation of knowledge such as facts, data, or opinions in any medium or form, including textual, numerical, graphic, cartographic, narrative, or audiovisual forms. A data sharing agreement may specify additional information that is to be provided between two or more parties.

## Metadata

From the NMFS Data and Information Policy Directive: Information describing the characteristics of data. Metadata can be data, information about data, descriptive information about an entity's data, data activities, systems, and holdings. Common uses for metadata include providing the context of the data resource, managing its lifecycle, and extending it to new uses. An example of metadata is the external description of an audio file specifying the artist that created it, when it was created, the length of playtime, and its genre of music. The purpose of metadata is to manage and improve the use of data and thereby turn it into a strategic asset.

## Open Data

Open Data is defined as structured data that is machine-readable, freely shared, used and built on without restrictions. The Open Data Act stipulated that the default state of new and modernized Government information resources shall be open and machine readable. Government information shall be managed as an asset throughout its life cycle to promote interoperability and openness, and, wherever possible and legally permissible, to ensure that data are released to the public in ways that make the data easy to find, accessible, and usable. In making this the new default state, executive departments and agencies (agencies) shall ensure that they safeguard individual privacy, confidentiality, and national security.

Ultimately "open data" speaks to a removal of conditions associated with the data. Once the data is "open," it can be used for anything and by anybody, so there is a release of control over the information by the persons or entities with control over the data. Given that NOAA must report on its obligations as a public interest regulator, there is a large subset of information that becomes public information as a result of the Open Data Act. As previously discussed, this openness can create additional uncertainty around what is "proprietary" or "confidential" information.

The openness of data is a newer lens to have to consider when assessing data sharing and data publishing operations. Open data can be used by virtually anyone. This makes it critical to ensure the information that accompanies the open data, describing the data methodology that was used for its creation is shared, and that there is capacity to support questions from the community of open data users. Open data is one of the most permissive arrangements for data use, it is an open-ended license, sometimes provided with a limited set of conditions.

## Public Information

Public information “means any information, regardless of form or format, that an agency discloses, disseminates, or makes available to the public” (per 44 U.S. Code Sec. 3502(12)). Information that is disclosed through the Open Data act becomes “public information.”

## Proprietary Information

In the context of fisheries data, two examples of proprietary information are fishermen's logbooks or their landing information, as both relate to their business practices and are information that they generally do not want shared with others. More generally, proprietary information is information that someone exerts a claim over, citing the exclusionary rights that are conferred through their ownership of the information or data. To minimize any risk of improperly sharing data, it is important to have explicit conversations about what any party to a data sharing agreement considers proprietary information, and why that information must be kept proprietary. Sometimes certain uses of proprietary information may be acceptable and others not. These things can only be discussed in context - assumptions about why certain information is considered important to keep proprietary should be interrogated to create a full understanding of the data being discussed.

### IT Systems & Technology Products

## Access to Data

Access to Data (also known as “data access”) is a long-standing IT practice that manages the way data is stored, used, and retrieved. While data access is a vital part of data management processes, and has a role in data sharing, the idea is more of an operational and internal practical matter than a foundation for data sharing agreements.

When data that should be shared internally appears to create a roadblock or a challenge, this may be more a matter of data access than data sharing, and may for that reason be managed more easily through a technical workflow (seeking database permissions, for examples) than the creation of a data sharing agreement.

## Application Programming Interface “API”

An API is how computers “talk” to each other, through one computer’s API which is designed to share data in a particular way, at particular times, in particular formats. Data may move through an API as part of a data-sharing agreement. This packaging of data presents one of several technical spaces that define how data is used. It may be used to keep track of changes to data services, and is an intentional mode of data sharing so the rules that define how an API works and who it is available to are pertinent to any data sharing discussions. Also pertinent is the identification of the stakeholders that will be in charge of the management of an API, to understand both how data will be used in an API and how changes to that method can be triggered and how they are managed.

## Artificial Intelligence

The term “artificial intelligence” is often used to describe one of the following: automation, machine learning, and general artificial intelligence. Advancements in and applications of artificial intelligence to fisheries management is an accelerant to digital transformation, one that prompts

the evolution of flexible, alternative data sharing options to better support robust data-sharing. A 2016 report from the White House, The Administration's Report on the Future of Artificial Intelligence, focused on the opportunities, considerations and challenges of Artificial Intelligence (AI). NOAA also has an Artificial Intelligence Strategy Fact Sheet.

## Automation

Automation refers to the techniques, methods, or systems of operating or controlling a process by automatic means, often through software, and with an intention to reduce human intervention to a minimum. Automation is a trend that is manifesting in the increasing digitization of fisheries management tools. This process tends to both increase the amount of data created and increase the need for new and complementary data-sharing tools that empower NOAA actors.

## Database

A database is an organized collection of structured information, or data, typically stored electronically in a computer system. Data-sharing agreements may grant mutual access to various proprietary databases.

## Data Brokerage

A "data brokerage" is a business that sells access to data sets. In the context of data sharing agreements, it is important to consider if a stakeholder of this type might be engaged at any point in a supply chain of data use and what the implications of that involvement would look like.

## Data Lifecycle Management

NMFSPD - A policy-based approach to managing the flow of an information system's data throughout that data's life cycle, from creation and initial storage to the time when the data become obsolete and are deleted. Forging a new data-sharing agreement or system necessitates consideration of the overall management of the data lifecycle.

## Digitization

"Digitization" refers to the conversion of text, pictures, or sound into a digital form that can be processed by a computer. Digitization of fishery data has created more complexity and prompted exploration of more flexible tools for improved data sharing.

## Information Lifecycle Management

Sometimes referred to as the data lifecycle, this process plans and manages the use and storage of data throughout its existence within an organization, from new additions to a data catalogue and the work required for labeling and maintenance all the way to archiving and potentially deleting data. This process is a helpful way to map data and its use over time when considering how a data sharing agreement should consider everything from information about the methodology used to create data to its maintenance and storage. The ongoing cost of supporting this process is high, and each step of its process should be considered when devising agreements for data use and sharing.

NMFSPD - A comprehensive approach to managing the flow of an information system's information and associated metadata through all stages: creation or collection through initial processing and storage, dissemination, and final disposition. ILM also includes user practices, planning, budgeting, manipulating, and controlling information throughout its life cycle.

## Internet of Things “IOT”

The “Internet of Things” refers to the concept of connecting physical objects (from fridges to cars to boats) to the internet via sensors. This increased connectivity between the physical world and the internet presents a range of new opportunities to track, monitor, measure, surveil, and control the physical environment. There is a lot of automated measuring and tracking that can occur through the IOT as well as a whole new host of vulnerabilities created by a persistent connection to the internet. The IOT is one of the transformational digital technologies that is prompting this exploration of more flexible direct data licensing to create new data-sharing agreements, and generally encompasses the technology that underpins Electronic Monitoring or Vessel Monitoring Systems (EM and VMS).

## Machine Learning

Machine learning describes software and algorithms that evolve and adapt based on the “learning” they do through the use of data. Machine learning is one of the applications of technology that has been consistently identified as being part of the digital transformation of fishery data management systems and applications. Machine learning is relevant to data and data sharing discussions because machine learning systems rely on large amounts of data to train their models and continually improve. As such, data is an input to these systems and how that data is used primarily, and then further along in the life cycle of machine learning, is worth noting as an object of value, a space of liability, and a commercial consideration for vendors seeking data use for machine learning products as well as researchers doing the same, among other purposes.

## Maintenance

NAO 216-100 - Maintenance is defined as the procedures required to keep confidential data secure from the time the source documents are received by NMFS to their ultimate disposition, regardless of format. Like many hard physical systems, data-sharing infrastructure and systems require constant care and if anything, even more maintenance than legacy tools. Our informational stakeholder interviews reflected a leadership divide in terms of the complexities and challenges inherent in the maintenance of the data system.

## Patent

A patent is a legal document which provides protection to the ideas of an individual. Usually issued by the Patent Office of a country, the patent is granted to any firm or individual. Patents generally fall into one of four different classes: Machine (a device or apparatus created by a person for the performance of a specific task, process (a process created by an individual), manufacture (any fabricated or manufactured product) or the composition of matter (any chemical mixture or compound created by a person). Patents enable parties to sue others who use their idea without getting the appropriate license.

## Rights Brokerage

A “rights brokerage” is an organization that sells the right to use data that has been aggregated and collected by the owner. In the context of data sharing agreements, it’s important to consider if a stakeholder of this type might be engaged at any point in a supply chain of data use and what the implications of that involvement would look like.

## Trade Secret(s)

“Trade secrets” include any valuable business information that derives its value from its secrecy. Historically, fisheries management has often treated fishing location as a trade secret and fishermen have preferred for catch location information to be shared only in aggregate.

In the context of data sharing agreements, data use, and technology, the idea of a trade secret may be defined as proprietary information about how data is used in the course of a commercial product such as software. In future conversations about data sharing and use, it will become increasingly important for regulators and those with a vested interest in providing accountability and transparency about data use to understand if and how trade secrets might be invoked by a vendor as rationale to keep data processing activities closed/confidential.

Trade secrets are also an important feature for any organization that deals with FOIA requests to understand, as public organizations may be held back from sharing information due to trade secrets/commercial information. This trend has been on the rise in recent years as governments have procured technology to establish algorithmic accountability standards and rules. One recent policy undertaking in New York City resulted in very little output due to the stakeholder group being unable to agree on what constitutes an “automated decision” and if and how the information behind an automated decision could be made known or available[25].

## Vessel Monitoring System “VMS”

A vessel monitoring system (VMS) consists of a NMFS -approved transmitter that automatically determines a vessel's position and transmits it to a communications service provider. Like a data-sharing agreement, VMS incorporate components of authority, consent, and data collection. VMS requirements vary by region and NMFS posts requirements online for: Alaska, Northeast (Greater Atlantic Region), Pacific Islands, Southeast, West Coast, and Atlantic Highly Migratory Species.

[25] “The First Effort to Regulate AI was a Spectacular Failure” Fast Company Magazine, Albert Fox Cahn, November 26, 2019.