Public fish – public data?

Kate Wing
the databranch
Oakland, CA, USA
kate@thedatabranch.org

ABSTRACT

The ocean is a vast commons, covering two-thirds of the globe. Ocean fisheries provide almost 60 million jobs and a critical food source for 3 billion people (FAO, 2018). Despite the ocean's importance, our data on ocean uses has historically been sparse and low-quality, particularly data for fisheries. The next five years will bring an inflection point for fish data, with the growth of satellite vessel tracking, on-board video cameras, and cheap sensors. Governments, fishers, scientists, and technology vendors are just beginning to explore the data and regulatory infrastructure needed to support this transition. This paper explores questions around accountability, ownership, privacy, and transparency where the ocean community seeks to learn best practices from other sectors as we develop solutions for the public's fish.

1.INTRODUCTION

The interconnected ecosystems of the sea require scientific and governmental cooperation beyond national borders. Large-scale research projects using satellites and oceanographic sensors already produce high volumes of so-called ambient monitoring data, such as ocean temperature. This type of passively collected data is generally shared across academic and government research institutions and can be used to make predictions that inform management decisions although it is not explicitly tied to policy goals. The other category of data – compliance data – tracks human activities and tends to be both scarce and siloed within agencies (Kelly, 2014). This can be the result of deliberate efforts to protect trade secrets or personal privacy, or it may be a legacy of old data practices. Whatever the reason, data scientists working on fisheries more often find themselves extrapolating from limited data rather than managing a flood.

When compliance data were expensive and difficult to produce, these gaps in our compliance data systems could be overlooked. Now, we are at an inflection point for ocean use data, as fishing boats carry detailed vessel tracking systems, countries require catch reporting via cell phone, and hand-held DNA sensors allow buyers and inspectors to verify seafood as it travels around the world. This provides an opportunity for the fisheries community to develop a policy framework for data collection, sharing, and stewardship before the wave of data arrives.

The public trust nature of the sea adds a layer of complexity to these policy discussions. Ocean resources are held in trust by governments and tribes, who are expected to act as stewards of long-term sustainability. The lack of private ownership in the ocean does not inure it from the private and commercial data enclosures seen in other sectors, but the stewardship obligations of the government may provide more options for public transparency and collaboration.

In this paper, I examine the issues of accountability, ownership, privacy, and transparency through the lens of video monitoring on U.S. fishing vessels. Video surveillance, also called electronic monitoring or EM, is expanding worldwide, with the potential to rapidly increase fisheries data by orders of magnitude, requiring new approaches to storage and automated analysis. If we can design systems that balance privacy and transparency for EM, if we can incentivize fishers' stewardship through creative data co-ownership, we can apply these methods to other fisheries data streams and ultimately provide better long-term stewardship for the ocean.

2.FISHERIES BACKGROUND

2.1.THE CHALLENGE OF COUNTING FISH

The core of modern fisheries management is the catch limit: the number of fish which can be removed each season while leaving enough fish in the ocean to reproduce and sustain future fisheries. Calculating catch limits requires fishery scientists to know or estimate multiple parameters which they use to model populations and predict future trends. Total fishing mortality is chief among these essential data elements, which means scientists and managers need to know how many fish are caught both to assess population status and to track how close they are to reaching that season's fishing target. Catch data serve both a science and a compliance goal.

Since commercial fishermen sell the catch they bring to shore, this landings data is the simplest way to estimate total catch and fishing mortality. Data from individual landing receipts is aggregated in state and regional databases, provided to scientists to model fish populations, and shared publicly in a highly aggregated form with time delays. For example, as of June 30, 2018 the latest publicly available national landings data on the NOAA Fisheries website was from 2016 and from 2017 on the Atlantic Coast Cooperative Statistics Program regional data portal.

However, not all fish caught are landed. Fish may be discarded at sea because they're low value, too small, or illegal to keep, and those discarded fish may or may not survive their return to the water. There are other at-sea activities managers need to track, like interactions with birds or marine mammals. To get that data, managers have required self-reporting via paper or electronic logbooks. In some fisheries trained observers ride on board a subset of fishing trips to report on catch and discards and take

Bloomberg Data for Good Ex
16-Sep-2018, New York City

scientific samples. Not all boats can safely accommodate another person, and there is a documented "observer effect" where boats fish differently when they carry an observer (Faunce and Barbeaux, 2011).

There are significant economic and ecological costs to getting catch limits wrong: overfish and populations can crash, close a fishery too early and fishermen and their communities lose income. There are other costs to relying on these legacy systems, from the high price of flying observers to remote ports to the wasted staff time when scientists spend a day entering data from paper logs. Fishermen, managers, scientists, and the public all have incentives to gather and use more accurate catch data, and the costs of legacy systems. Enter electronic monitoring.

2.2.ELECTRONIC MONITORING

Electronic monitoring, or EM, uses a network of on-board sensors to track vessel location and fishing activity. These generally include cameras in addition to sensors on gear and hydraulics that detect when gear has been deployed, such as unspooling a net or throwing a trap overboard. EM systems run continuously, from when a boat leaves the dock to when it returns, and some percentage of that footage is reviewed to identify catch, discards, and protected species interactions. This lets EM avoid the "observer effect" as every trip is recorded, and can be configured to cover almost any boat size or gear type, so it can operate on boats too small for human observers.

All current EM programs involve partnerships across government, the fishing industry, and EM vendors. The main ingredients of an EM system are the on-board hardware, data transmission, data review and storage, any one of which could be assigned to government, industry, a vendor, or a mix of the three. For example, fishermen could build their own hardware with off-the-shelf electronics and hire a vendor to review the video, or government could own the hardware and move it from boat to boat. The actual modes for transmitting data also vary, with some vendors recording video on physical hard drives that fishermen mail in when they return from a trip, while others process video on-board and return data via satellite. Reviewing video to count and identify fish happens on land, with vendors storing the raw video and providing tabular data to government. Because the EM market is relatively young, the handful of existing EM vendors are still experimenting with their products and services. This means the available options, including data licensing arrangements, are highly sensitive to government rules and requirements as vendors try and anticipate customer needs.

Groundfish boats in British Columbia have carried EM systems since 2006, but until recently there have only been limited pilots in the United States. Now, federal managers have indicated they want to expand EM in the next two years, and some states are also exploring putting cameras on state-managed commercial fishing boats. This would mean expanding from fewer than 50 boats to more than 500 in 2020 and upwards of 5,000 boats in 2025. This is essentially a tsunami of petabytes of video data headed towards a data system and a policy framework built in the 1980s. The coming wave has prompted people from across the fisheries community to talk about data stewardship, automation, and infrastructure in the fisheries sector.

3.KEY ISSUES

3.1.DATA OWNERSHIP AND STEWARDSHIP

In general, U.S. fishermen have been able to keep copies of data they generate (e.g. landing receipts and logbooks) but not necessarily data recorded about them (e.g. observer reports). Multi-party licensing arrangements from EM vendors change the options available to fishermen, and they require government to provide clear guidance about data control and access. For example, an agency could sole-source a contract directly with one vendor, require fishermen to use only that vendor, and make itself the sole data receiver. While this may satisfy compliance and chain of custody concerns, it's problematic for other reasons, including making government assume all the costs.

The cost issue is one reason agencies are exploring making fishermen the owners of the raw data, while government takes responsibility only for submitted data, which becomes a government record upon submission and may include only tabular data from a reviewer, not images. Vendors would be responsible for ensuring tamper-proof systems, and a clean chain of custody for data submissions, but fishermen could access, download, and reuse raw video and extracted data as they wish.

If government decides to rest data ownership with the fishermen, this opens a large window for new data products and opportunities. Vendors could provide business analytics to vessels and fleet managers, fishermen could attach images to shipments to visually document a fish's journey from boat to plate, or fishing associations could form data trusts where members share and sell their data. The Cape Cod Commercial Fishermen's Association is working with Keith Porcaro at Digital Public to explore a digital trust, with one potential use being documenting safe fishing practices for insurance companies.

Conversely, if the government takes only a limited amount of extracted data, relinquishes most access to the raw footage, and requires very short retention times, they may not be meeting their public transparency responsibilities. We're only beginning to explore the scientific value of onboard video footage, and if the data disappear or are publicly inaccessible at all but the grossest level, we may be missing an opportunity to improve stewardship and hold the agencies to task for their public trust mandates.

3.2.ALGORITHMIC ACCOUNTABILITY

Automation is essential for broader EM coverage. Paying people to sit in windowless rooms and stare at hours of footage is cost prohibitive and not the best use of fisheries expertise. Some EM vendors already apply AI in their workflow: Integrated Monitoring transmits one frame every six seconds via satellite until its AI detects motion in a camera area, prompting the system to increase transmission. Other AI projects have trained algorithms to count fish, identify species, and measure their length. One of these efforts, which I was involved with, deliberately made the algorithms and video open both so that anyone could add them to their EM workflow and because we wanted them to be auditable.

There is a growing interest in creating a library of open, labeled imagery that would support the development of new fish AI to help automate review and reduce costs, as well as allow some testing and review of proprietary vendor products. Getting that imagery requires fishermen-owners to share it and/or procedures that allow governments to release video edited to protect privacy. Per the previous discussion of data ownership and stewardship, purchasing video from fishermen's associations to make training and testing datasets could be one way to reward early adopters of EM.

The impact of faulty AI to control video transmission rates is low, as long as the full recording still exists. It becomes greater when AI is used to deduct fish from individual catch accounts, penalize fishermen for violations, and add data to biological assessments affecting the whole fishery. At a larger scale, three projects (Skylight, Global Fishing Watch, and OceanMind) are feeding their AI multiple data streams to find and predict illegal fishing on the high seas. We need to continue to push for transparent operation and results in fisheries AI to make sure we're getting accurate results.

There's also an employment impact of greater automation, as AI takes on tasks previously done by human reviewers, often former observers. In remote areas of the U.S., as well as in other fishing nations, these are difficult but good jobs. Even the best-performing algorithms will require expert human assistance to interpret difficult images, so agencies should prepare for a human-in-the-loop AI approach. Not all observers will want to shift from an outdoor professions to watching a screen, which means some will leave an EM-based workforce. Conversely, video review may attract more women than the heavily male-dominated field of on-board fisheries observers.

3.3. Privacy & Confidentiality

The law governing U.S. fisheries adds protections to fishing data on top of protections provided by the Privacy Act and Freedom of Information Act (FOIA):

> "…the Secretary may release or make public any such information in any aggregate or summary form which does not directly or indirectly disclose the identity or business of any person who submits such information." (16 USC 1881§b(3))

When Congress passed this language in 1976, they did not envision a time when anyone could track the position of boats in real time or view high-resolution satellite imagery showing gear in the water. NOAA Fisheries, the federal agency charged with U.S. fisheries management, has interpreted the charge to protect the "identity or business" behind the data using a type of Heisenberg Uncertainty principle. The public can know who owns a vessel, but not what it has caught, when, or where. Fishing location is seen as a trade secret and given the highest level of protection. Even within the agency, data for individual vessels or trips is tightly controlled between offices and staff.

While NOAA Fisheries is taking great pains to conceal the identity and businesses of fishermen, particularly fishing locations, that data can now be publicly reconstructed. Fishermen themselves are posting photos of their catch on Facebook and Instagram, including video from their EM feeds, and much of that data includes geotags. Even if agencies try to strictly control access to EM data that is a government record, they have limited powers to prevent individuals from sharing their own files.

Depending on what EM data NOAA Fisheries decides to store, past approaches to responding to FOIA requests may no longer apply. While the agency primarily require tabular data submissions, they are likely to still receive and retain some imagery, particularly in cases where they need to prosecute violations. In one example where observer photos were FOIA-ed, the agency printed and mailed the images, in part to avoid releasing metadata that could have revealed observer or vessel identity. This is unlikely to be a workable solution at scale. In states with Public Records Acts, data sharing protocols that worked to protect tabular data (e.g. landings, permit-holder identity) need to be upgraded to handle EM data.

3.4. Transparency and Data Literacy

In the United States, eight regional councils manage federal fisheries. Council members include state and federal agency staff along with members appointed by Governors, usually fishermen and industry representatives. The Councils meet publicly and take advice from their scientific and statistical committees in setting catch limits and other management measures. While this provides a public forum to influence and review fisheries decisions, it has not often provided a way to introduce new data, due to concerns about bias and poor data quality from outside sources. In the future, fishermen could pool their EM data to run their own discard and catch analyses using the same dataset as the agencies, and use that information to challenge council decisions. This could open the door to accepting more public science as well, as apps and sensors can verify and authenticate data. We need to revisit protocols for accepting data into the management processes to make sure we are not inappropriately excluding or including new data sources.

Finally, the overall level of data literacy in the fisheries community needs to increase. Fishermen and agency staff may not realize how much private data is contained in image metadata. Blurring faces and vessel IDs in a video clip may not be sufficient to protect personal or business privacy. Crew and captains should be actively informed about how footage of them will be used and shared. Management council members will need to be ready to question data scientists and more complex models.

4. CONCLUSION

The next five years offer an opportunity for managers, fishers, scientists, and the public to participate in designing frameworks for data that improve our understanding of the ocean and incentivize long-term fisheries sustainability. While the ocean's complex governance structure and salty nature provide some unique monitoring challenges, the issues of balancing privacy and openness, of transparency and trust, are issues shared by many fields. If data are a valuable resource, then those who steward the ocean well should share in the value of that data, as they share in the benefits of well-managed fisheries. Our shift to a data-rich future should provide the public a window into the public sea.

5. ACKNOWLEDGEMENTS

Work on this paper was supported by a fellowship from The Nature Conservancy. Thanks to John Wilbanks and Lucy Bernholz for inspiring me to look in unexpected places for ocean inspiration.

6. REFERENCES

FAO. 2018. The State of World Fisheries and Aquaculture 2018 - Meeting the sustainable development goals. Rome. Licence: CC BY-NC-SA 3.0 IGO.

Faunce and Barbeaux, 2011. "The frequency and quantity of Alaskan groundfish catcher-vessel landings made with and without an observer" *ICES Journal of Marine Science* 68(8)

Kelly, R.P, 2014. "Will More, Better, Cheaper, and Faster Monitoring Improve Environmental Management?" *Environmental Law* 44, 1111