Planning for a New England Groundfish Data Portal

December 2021

"We're entering a new world in which data may be more important than software."
– Tim O'Reilly, founder, O'Reilly Media.

Contents

Summary

The future of small-boat commercial fishing communities around the country depends on confronting and solving the challenges of fishing accountability and data modernization. Timely, accurate, unbiased, and trusted data collected during commercial fishing trips (fishery-dependent data) is essential to rebuilding fish stocks using the best available science. Fishermen's access to their own fishery-dependent data is also important, for supporting business decisions, understanding impacts of warming waters, verifying accuracy of data, and building trust in the management process.

Recent endeavors in New England fisheries have set the stage for improving data access and maximizing the value of data collected on commercial fishing vessels. These include NOAA Fisheries' data modernization efforts and the implementation of electronic monitoring (EM) programs with the associated creation of novel EM data streams. It is time to create a pathway for the New England fishing community to access data generated on their vessels, through the creation of a regional, cross-fisheries, community-centered data platform that provides fishermen with business decision making tools from their fishery dependent data while supporting independent science and improved management.

This report details the results of an 18 month effort to document available data and design a New England Groundfish Data Portal, with the primary goal of accommodating the needs of commercial fishermen. Secondary goals included supporting the needs of non-NOAA scientists and groundfish sector managers while building a strong foundation of relationships, solutions and tools, and commitments to change. In the process of designing the data portal, a prototype was developed and piloted with significant documentation and code, to support future efforts to build out a functioning data portal and user interface. Industry user case scenarios were developed to support application design and functionality.

The resources developed through this process are freely and publically available, with the goal that they will be used by one or more of the following entities to improve regional data access:

1. NOAA Fisheries Greater Atlantic Regional Fisheries Office and Northeast Fisheries Science Center use the guidance to design improved industry data access and a fishermen-friendly user interface. For instance, they could integrate the portal design into Fish Online and their new Catch Accounting Management System (CAMS), to dramatically improve fishermen's access to their own data and provide easy to use visualization and analyses tools. This option has no data sharing or data access barriers, since NOAA already stores all of the data and recently improved the industry's Fish Online access with secure online access to selected fisheries datasets.
2. An independent, neutral third party builds and hosts a community data portal to support fishermen's access to data. This option will require data sharing agreements and negotiating data access with NOAA Fisheries or Sector Managers. It has the largest hurdle in terms of ongoing operating costs. Unless

real time individual quota balances are available to fishermen, they are not willing to pay modest annual or monthly fees for data portal services.

3. Electronic Monitoring Service Providers incorporate all or portions of the data portal design to provide value-added services to their fishermen clients through online access to data and visualization tools. This could be a free service, to differentiate a service provider's offerings and make it more competitive, or it could be provided for a small additional charge that gets wrapped into the EM price package. There are no data access barriers to providing electronic monitoring data to the fishermen, besides maintaining confidentiality. Integrating eVTR and dealer data streams would require data sharing agreements and negotiating data access with NOAA Fisheries or Sector Managers.

Background

The future of small-boat commercial fishing communities around the country depends on confronting and solving the challenges of fishing accountability and data modernization. As a region, New England is poised to be a national leader in catalyzing changes in both fisheries monitoring and data management. To do so, we must:

- communicate the best way to rebuild groundfish stocks and fishing businesses, which is to ensure full accountability in the fishery with strategies like electronic monitoring, and
- ensure that the data streams and science used for stock assessments is timely, accurate, unbiased, and trusted by both the scientific and management agencies that rely on the data to make decisions and the fishing communities that are impacted by those decisions.

Meaningful improvements to fishery accountability can maintain livelihoods today while rebuilding stocks and sustainable ocean ecosystems for tomorrow. Opportunities can be created for fishermen to see accountability as an investment that rebuilds fish stocks, providing increases in quota.

If New England can achieve accountability, modernize data management, and align on-the-water reality with science and management, it will result in improved fisheries management decisions, more economically efficient quota utilization, and a sustainable, recovered fleet. For instance, interoperable data systems and unique trip identifiers make analyses of fisheries-dependent data more efficient and expeditious for both government and non-government users, as well as providing the foundation for seafood traceability. Without an integrated approach, accountability alone will likely consolidate quota on corporate vessels and destroy what remains of New England's historic, small-boat fishing communities. A thoughtful transition is of greater importance to small businesses than to vertically integrated corporations. Outside of the fishing communities, an integrated approach that embraces and takes advantage of data modernization will provide New England the tools and data that are necessary to respond to broader ecosystem challenges, and serve as a model for other fisheries and regions.

The Cape Cod Commercial Fishermen's Alliance received grant funding from the Net Gains Alliance to develop plans for a regional fisheries data access portal, including working with fishermen and interested scientists to define end-user needs, and working with NOAA to ensure legality of data access and technical specifications. Ultimately, the goal is to ensure fishermen can access the value of their fishing data and are empowered to use it. This work builds off of preliminary efforts in 2018-2019 to explore opportunities and models for the groundfish community to own, control, manage, and use fisheries data (see Appendix 5: Digital Public Report: Building a fishermen-first data ecosystem)

This portal, once built, will dramatically accelerate the timeline for providing data users with legal, value-added, and timely access to electronic monitoring (EM) and other fishery-dependent datasets. It will also provide fishermen with secure access to their own confidential data in an easy-to-use format that provides value-added visualizations and analysis tools for business planning, incorporating other regional publically accessible data sources like water temperature, wave height, and tides. Fishermen's data access is an important step in providing fishermen the resources they need to efficiently and effectively manage their fishing businesses and ensure that small-boat fishing communities survive and thrive.

In the process of designing a fishermen-centric data portal, we focused on the New England groundfish fishery, given its data-heavy management and accountability requirements. As a catch share fishery, it is managed in-season with weekly quota reports and starting in May 2022, will require at least 99% at sea monitoring. It is operationalizing electronic monitoring programs and mandating electronic reporting. It is poised to serve as a national example of how to access, integrate, and utilize fishery-dependent data.

This report documents the process and findings of an 18 month effort to create a pathway for the New England fishing community to access and utilize data generated on their vessels. The design process incorporated information and opinions from seasonal groundfish fishermen, groundfish Sector Managers, EM Service Providers, regional non-profits Gulf of Maine Research Institute (GMRI), The Nature Conservancy (TNC), and Responsible Offshore Development Alliance (RODA), NOAA Fisheries staff from the Greater Atlantic Regional Fisheries Office (GARFO) and Northeast Fisheries Science Center (NEFSC), the eMOLT Program, the Northeastern Regional Association of Coastal Ocean Observing Systems (NERACOOS), and academic, non-profit, state, and federal scientists.

Methodology

The design project was launched in June 2020 with general outreach and introduction of goals to NOAA Fisheries staff, academic, non-profit, state, and federal scientists, and commercial groundfish fishermen. Given the COVID-19 pandemic, all of the planned in-person meetings and workshops were converted to online and electronic formats.

When originally approached about this project, NOAA Fisheries GARFO and NEFSC leadership committed staff time to participating in the portal design. As 2020 and 2021 progressed with continued COVID work from home restrictions, reduced staff capacity, increased commitments and deadlines to advance electronic monitoring/electronic reporting and data modernization, and new COVID-related protocols for the fishery observer program, neither GARFO nor NEFSC was ultimately able to contribute meaningful staff time to participate in the project. They were able to answer several specific questions around data formatting and data management. Project staff were able to leverage several Electronic Monitoring Program meetings to hold high level discussions with NOAA staff regarding the importance of fishermen data access, unique trip identifiers, and modernizing NOAA's data management.

The majority of effort was focused on fishing industry engagement, to document commercial groundfish fishermen's data needs. Discussions were also held with groundfish Sector Managers and EM Service Providers. Feedback and discussions occurred with more than 26 unique industry individuals through the following mechanisms (several participated multiple times):

1. Electronic Survey to gauge data priorities and interest in data sharing
2. Three Workshops with short visual presentation of examples and in-depth discussions on portal data sets and functionality.
3. Functioning online Data Portal Prototype to test, loaded with mock data (available online for four months)
4. Electronic Survey to gauge response to Data Portal Prototype
5. Individual conversations between project staff and end users
6. In-depth pilot testing of a secure version of the Data Portal Prototype, loaded with their own real, confidential data and detailed responses

Industry guidance throughout the project was used to develop and refine the Data Portal Prototype and to develop six user case scenarios, detailed in Appendix 1: Industry User Case Scenarios. The project relied upon staff from the Fishermen's Alliance and GMRI to engage with industry members, conduct workshops, and document findings. Danielle Quinn, an ecoinformatics data science consultant, was contracted to organize the datasets, design code and analytics for the data portal prototype, pilot the prototype with fishermen, refine the prototype, and provide technical documentation. Mock data was created from using data file templates from real groundfish data provided from GARFO, EM Service Providers, and the Sector Manager's SIMM data downloads; this ensured proper formatting and use of available data fields while providing industry with realistic data examples to review in the prototype. Several fishermen commented that they would be able to provide more meaningful feedback if they could pilot the prototype with their own, real data. As a result, selected fishermen received secure, personal links to their unique version of the data portal prototype, loaded with their actual fishing data from FY2019 and FY2020. Their extensive feedback was incorporated into the user case scenarios and technical documentation.

Academic, non-profit, state, and federal scientists throughout New England were sent an online survey to gauge on their interest in and use of fishing data, with particular

focus on the novel electronic monitoring (EM) data stream. Many individuals represented larger institutions or lab groups that could make extensive use of these data. The respondents come from a range of sectors including non-profit research organizations, government agencies, academia, and environmental consultants. They completed the multiple choice survey and provided written text responses. The results are detailed in Appendix 2: Scientist User Case Scenario.

It is important to note that ultimately the vast majority of fishermen (92%) were not interested in aggregating their data at a fleet wide level to share with scientists nor other fishermen, even if confidentiality was maintained. 42% of fishermen were amenable to sharing their data with their sector manager or a specified individual. As a result, scientist access was not built into the data portal prototype.

The project discussed merging efforts and data portals with the Responsible Offshore Development Alliance (RODA), since they are also in the process of aggregating fishing data into a portal to support decision making and science around offshore windfarms. It was ultimately decided that there was not enough overlap in project goals and data resolutions to warrant merging efforts.

The Northeastern Regional Association of Coastal Ocean Observing Systems (NERACOOS) provides robust environmental data sets that are relatively easy to access and incorporate through the Environmental Research Division's Data Access Program (ERDDAP) online data portal. This included water temperature from an array of ocean buoys in the New England region. Details on how this was accomplished is included in the technical documentation. NERACOOS staff is supportive of the groundfish data portal project and is ready to assist as needed to provide data sets during full development.

A demonstration of the Data Portal Prototype was provided to NOAA Fisheries' eMOLT Program and their technical staff helped troubleshoot our efforts to properly display and incorporate eMOLT bottom temperature into the data portal. eMOLT is also in the process of building out online data access for eMOLT fishing vessels and may utilize portions of the Data Portal Prototype. While their currently funded scope is only to return the eMOLT data to the fishermen, they are interested long-term in how to incorporate catch and effort data. If the Data Portal is fully developed, it may negate the need for the eMOLT program to build its own version.

It should be noted that since the third party, independent data portal project was conceptualized in response to NOAA's poor data access, NOAA Fisheries in New England has made some significant advances in and plans for improved data management and access, as outlined in its August 2021 Electronic Technologies Implementation Plan; it is on the cusp of taking delivery on more than a decade of work to modernize data infrastructure and management. These changes create stronger support for NOAA Fisheries to provide a user interface like the Data Portal for fishermen to access their data, instead of an expensive third party with complex data sharing agreements. Examples of NOAA's progress include:

- The Fishery-Dependent Data Initiative (FDDI) led by GARFO and NEFSC may provide useful solutions including the use of a Universal Trip Identifier (UTID) and the Catch Accounting Monitoring System (CAMS), which in combination could streamline data processing and integration of multiple and duplicative data sources into a single database that is shared by GARFO and NEFSC. It is anticipated that the CAMS system will be operational in 2022.
- GARFO upgraded their Fish Online system and fishermen can now log in to submit eVTRs, review landings and days at sea data, and manage permits. The upgraded secure log-in infrastructure allows multiple vessels to be linked to a single account.
- In the summer of 2021, GARFO moved the processing and management of eVTRs from an in-house server (that had a lot of challenges) to a cloud-based service. This change has dramatically increased the up-time of the VERS database and has allowed for near real time tracking of eVTR submissions by fishermen or sector managers (previously there was a 24-48 hr delay after submission). GARFO staff have indicated that the groundfish sector SIMM data server will also be moved to the cloud in the near future.
- The development of electronic monitoring programs has advanced NEFSC's data management and introduced an EM API, found on the Northeast Electronic Monitoring Information System website (NEMIS). This API technology is key for a streamlined and efficient data portal.
- Mandatory electronic vessel trip reports went into effect in November 2021. These eVTRs enable, for the first time in history, a unique trip identifier to link all the different fishery dependent data streams. NOAA is working to consolidate and integrate reporting requirements into the eVTR application to streamline reporting, reduce redundancies and improve operational efficiency.

Given these advances in data modernization, the most logical path forward for the Groundfish Data Portal is for NOAA Fisheries to adapt the proposed data portal interface into its data systems and provide fishermen with direct access to their data and visualization tools.

Data Portal Blueprint

Significant documentation was created for supporting the final development of the Data Portal, including data standards, metadata, R code, example data, and other supporting files.

These resources are available online under a GNU General Public License v3.0. at https://github.com/CCCFA/GroundfishDataPortal. All code, example data, and other processes are located in the "clean-dash" folder.

The Data Portal Developer Guide is a technical document that details data requirements, data pre-processing, reference tables, R packages, environmental data, data portal structure and processes for workspace and data preparation, user interface,

and reactive server. Additionally, it documents user testing/feedback, next steps and wishlists for additional functionality. The Guide can be reviewed at https://github.com/CCCFA/GroundfishDataPortal

Please note that the example data available in the Prototype and in the GitHub repository is mock data based on data structure that is current as of 2021. They are only for demonstration purposes and should not be used for decision making or final data portal development. Updated data files should be secured based on final data access/data sharing agreements

The Data Portal Prototype, loaded with mock data, can be accessed online at: https://daniellequinn.shinyapps.io/clean-dash/.

Additional metadata and reference tables to support the data used in the portal include:

- Species table to standardize the reporting of species across data sources to the standardized names (AFS).
- Over the course of the project, we developed a metadata catalogue describing many of the data sources collected by the groundfish fleet in New England. Data sources documented in the catalogue include:
    - electronic vessel trip report (subset of all collected data fields) and dealer data available to sector managers through the SIMM portal.
    - electronic vessel trip report available to NOAA (broader)
    - universal standard EM summary report (what EM service providers submit to NOAA via the API)
    - EM discard file available to sector managers through the SIMM portal
    - at sea monitor trip reports available to fishermen and sector managers
    - eMOLT bottom temperature dataset available through the NERACOOS servers.

  The data catalogue was built with the online platform Datalogz.io but is no longer available online. A copy of the Metadata Catalogue is available in https://github.com/CCCFA/GroundfishDataPortal
  For each data source, the catalogue includes the following:
    - Dataset name
        - Storage location / point of access
        - Brief dataset description
    - List of data fields including:
        - Name
        - Data type (string, integer, decimal, date, etc.)
        - Description (formatting, standards, uniqueness, etc.)

Documentation detailing user preferences and priorities for the data portal can be found in the Appendices:
Appendix 1: Industry User Case Scenarios
Appendix 2: Scientist User Case Scenario

Data Sharing Policies

Based on fishermen feedback collected during the project, the following Data Portal guidance for data sharing policies and user account access has been developed. As noted above, the majority of fishermen are not interested in sharing their data broadly in the data portal with other sector members, academia, or the fleet. It is not recommended that the portal be built with the purpose of aggregating and sharing data among a larger group of fishermen. The focus should be on the owner, captain and sector manager level. For the purposes of documenting findings, these broader sharing levels were created using feedback from the few fishermen that were interested in sharing data.

Account Levels:

- **Owner:** The owner of the fishing vessel has access to 100% of the data collected on their fishing vessel.
  - Access all vessel-specific data, including the vessel's eMOLT data if applicable, at the temporal and spatial resolution it was collected at
  - Access all regional environmental data
  - Ability to download summary reports and raw data
  - Can assign user account access to other individuals, for them to view the data, permissions based on their account level
  - Can access multiple vessels' data and view reports, data and analyses at the individual level, aggregate level (all vessels together), and sub-aggregate level (selected vessels together). Vessels based on those owned by Owner.
- **Captain:** In cases where the captain is not the owner of the vessel, this user level has similar but not identical permissions as the owner.
  - Access all vessel-specific data for trips where captain (VTR) = **Captain**, including the vessel's eMOLT data if applicable, at the temporal and spatial resolution it was collected at
  - Access all regional environmental data
  - Ability to download summary reports and raw data for trips where captain (VTR) = **Captain**
  - Depending on owner preferences, the Captain may be assigned additional vessel level permissions (access to data collected on said vessel by another captain).
  - No ability to assign user account access to other individuals
- **Sector Manager:**
  - Access all vessel-specific data, including the vessel's eMOLT data if applicable, at the temporal resolution it was collected at
    - spatial resolution reduced to rounded decimal degrees DD.DD
  - Access all regional environmental data
  - Ability to download summary reports and raw data
  - Can access multiple vessels' data and view reports, data and analyses at the individual level, aggregate level (all vessels together), and sub-

aggregate level (selected vessels together). Vessels based on those assigned to the Sector Manager.
    - No ability to assign user account access to other individuals
- **All Access Permission:** In cases where the owner wants to give a specified individual access to data collected on the owner's vessel
    - Access to specified vessel-specific data sets at the temporal and spatial resolution it was collected at. Owner to specify which datasets are assigned to this user account.
    - Access all regional environmental data
    - Owner to specify permission to download summary reports
    - Owner to specify permission to download raw data
    - Can access multiple vessels' data and view reports, data and analyses at the individual level, aggregate level (all vessels together), and sub-aggregate level (selected vessels together). Vessels based on those assigned to the same All Access Permission account (each Captain would have to assign permissions to the same account).
    - No ability to assign user account access to other individuals
- **Sector Level Sharing:** Requires at least 3 vessels owned by separate entities (self plus 2 others) to opt into Sector Level Sharing with each other.
    - Access subtrip level data that has been aggregated across 3 or more vessels, with all location data adjusted to:
        - spatial resolution reduced to rounded decimal degrees DD.D
        - temporal resolution reduced to date (no time of day)
        - remove dealer ID field
    - Access all regional environmental data
    - Ability to download aggregated summary reports
    - No ability to download raw data
    - Can access multiple vessels' data and view reports, data and analyses at the aggregate level (all vessels together). Vessels based on those assigned to the same Sector Level Sharing account
    - No ability to assign user account access to other individuals
- **Fleet Wide Sharing:** Requires at least 3 vessels owned by separate entities (self plus 2 others) to opt into Fleet Wide Sharing.
    - Access trip level data that has been aggregated across 3 or more vessels, with all data adjusted to:
        - spatial resolution reduced to 10 minutes squares
        - temporal resolution aggregated to week
        - remove dealer ID field
        - eMOLT data aggregated to depth contour, daily average for that depth
    - Access all regional environmental data
    - Ability to download aggregated summary reports
    - No ability to download raw data
    - Can access multiple vessels' data and view reports, data and analyses at the aggregate level (all vessels together). Vessels based on those assigned to the Fleet Wide Sharing.

- No ability to assign user account access to other individuals

> **Scientific Analyses:** Requires at least 3 vessels owned by separate entities (self plus 2 others) to opt into the same Scientific Analyses account.
  - Access haul level data that has been aggregated across 3 or more vessels, with all data adjusted to:
    - spatial resolution reduced to rounded decimal degrees DD.DD
    - temporal resolution reduced to hours (remove minutes and seconds)
    - remove dealer ID field
    - if no haul level data available for a vessel, use subtrip level data.
      - Note, at least three vessels need to have same level or use the lower resolution for all (need at least 3 vessels with haul level to display haul level)
    - Access all eMOLT data, rounded to decimal degrees DD.DD
    - Remove gear type and related fields if not at least three vessels of each gear type
  - Access all regional environmental data
  - Ability to download aggregated summary reports and adjusted raw data
  - Can access multiple vessels' data and view reports, data and analyses at the aggregate level (all vessels together), and sub-aggregate level (selected vessels together, at least three). Vessels based on those assigned to the Scientific Analyses account.
  - No ability to assign user account access to other individuals

Future Plans for Data Portal Development

The resources developed through this design process are freely and publically available online (see Data Portal Blueprint section) with the goal that they will be used by one or more of the following entities to improve regional data access:

> NOAA Fisheries Greater Atlantic Regional Fisheries Office and Northeast Fisheries Science Center use the guidance to design and improve industry data access. For instance, they could integrate the portal design into their new Catch Accounting & Monitoring System (CAMS) and Fish Online, to dramatically improve fishermen's access to their own data and provide easy to use visualization and analyses tools. This option has no data sharing or data access barriers, since NOAA already stores all of the datasets. Improving data access and visualization tools is a priority of the Regional Office, but they have not yet had the resources to execute. The data portal design work should dramatically streamline the process of realizing an online data interface, as the prototype code is already written to integrate the data streams according to the Regional Office's metadata, overcoming the historic challenge of no unique trip identifiers.

> An independent, neutral third party builds and hosts a community data portal to support fishermen's access to data. This option will require data sharing

agreements and negotiating data access with NOAA Fisheries or Sector Managers. It has the largest hurdle in terms of ongoing operating costs. Unless real time individual quota balances are available to fishermen, they are not willing to pay modest annual or monthly fees for data portal services. In the process of designing the data portal, only one entity expressed real interest in building and hosting the data portal: Gulf of Maine Research Institute (GMRI). As a neutral third party in New England fisheries, GMRI could be an ideal host, if enforceable data sharing agreements with the industry ensure that GMRI cannot use the fishery-dependent data for its own research purposes without explicit permissions from the fishermen.

- Electronic Monitoring Service Providers incorporate all or portions of the data portal design to provide value-added services to their fishermen clients through online access to data and visualization tools. This could be a free service, to differentiate a service provider's offerings and make it more competitive, or it could be provided for a small additional charge that gets wrapped into the EM price package. There are no data access barriers to an EM service provider providing electronic monitoring data to the fishermen, besides maintaining confidentiality. Integrating eVTR and dealer data streams would require data sharing agreements and negotiating delayed data access with NOAA Fisheries in a way that ensures the EM Service Provider cannot access or use this data prior to reviewing video and creating the EM data stream.

The success of a data portal depends on reliable data access, especially if it is built by a non-NOAA, third party entity. Access to fishery-dependent data may be achieved in several ways, across different timelines, and will likely be a combination of two or more options:

- NOAA Fisheries allows a direct, automated connection to the data sets through an open API, once the appropriate data confidentiality protections and data sharing agreements are in place. This would provide the closest to real time data, be most efficient, and allow for the most automation of data access. There is precedent for this in the Northeast Electronic Monitoring Information System, which is how EM data flows from the third party EM Service Provider to the Northeast Fisheries Science Center. See the documentation tab at: https://apps-nefsc.fisheries.noaa.gov/NEMIS/
- NOAA Fisheries provides a regular data releases via Kiteworks data transfer to the portal on a regular basis (monthly, quarterly, or annually), once the appropriate data confidentiality protections and data sharing agreements are in place. The data portal manager will need to access the files on Kiteworks, download, and transfer into the data portal's secure database. The formatting of these data files must be standardized and identical for each transfer to ensure that the portal can read the data.
- NOAA Fisheries creates a SIMM account for the data portal, which allows the data portal manager to manually log in and download the same data files that sector managers access to manage the sector quota (eVTR, EM discards, dealer landings). This will also require that the appropriate data confidentiality

protections and data sharing agreements be in place. Occasionally the SIMM data fields change, which will require updates to the data portal code.

- EM Service Providers provide regular data releases of the EM summary data to the data portal, likely as a regular secure file transfer on a weekly or monthly basis, once the appropriate data confidentiality protections and data sharing agreements are in place. This could potentially be semi-automated if a secure, one-way, online dropbox is created where the EM Service Provider uploads the file and the portal can automatically read new files and incorporate into the data portal secure database.
- Sector Managers agree to download data from their SIMM account and release it to the data portal, likely as a regular secure file transfer on a weekly or monthly basis, once the appropriate data confidentiality protections and data sharing agreements are in place. This could potentially be semi-automated if a secure, one-way, online dropbox is created where sector managers upload the file and the portal can automatically read new files and incorporate into the data portal secure database.
- Fishermen access their confidential data from various sources and upload it the portal. This is not ideal for many reasons and is not recommended. Fishermen could get their eVTR, EM discard, and dealer landings data from their sector manager (likely at an additional charge for processing the data) and upload it to a secure, one-way, online dropbox. There is plenty of room for error here, in field formatting (trip ID numbers are often lost to scientific notation if the CSV file is saved as an excel file without changing the field to the correct number format), missing data, duplicative data, and user frustration. Fishermen may be able to access and download portions of their data from Fish Online. It may be feasible for data sharing agreements to be created similar to Sea State Inc.'s Bering Sea Reporting, where the fishermen share their Fish Online passwords and the data portal manager logs in on their behalf to download and scrape their individual data into the secure data portal database.

Note, there are other data sets that fishermen are interested in that we could not successfully access during the data portal design phase. To provide the most value to fishermen, additional work should be completed to access and assemble the following data sets:

- Ex-Vessel Prices for each species and market class. Ideally, this is the actual price paid to the fishermen on each trip, as reported by the dealers to NOAA. Fishermen cannot currently access this data except by looking at their weekly paper settlement checks from the dealers. Even then it may not be clear the exact price per pound. At a bare minimum, NOAA should aggregate the prices provided by the dealers and provide a weekly regional average ex-vessel price for each species and market class.
- Quota Lease Prices for each stock. Fishermen would like to track what they are paying for quota lease prices in the portal, to help calculate the cost and profit of their catch per unit effort (CPUE). Additionally, they would like to access a weekly

regional average quota lease price for each stock to compare to what they are paying. This could be created by NOAA from the weekly lease transaction reports submitted by the Sector Managers.

Funding Resources

Creation of a functioning data portal aligns with the priorities of several philanthropic grant or federal award programs. Funding for building the portal could potentially come from one or more of the following sources:

- National Fish & Wildlife Foundation Fisheries Innovation Fund. They seek to fund projects that partner with fishermen, stakeholders, state agencies and Fishery Information Networks to systematically integrate technology into fisheries data collection and observations, and streamline data management and use for fisheries management. Projects should catalyze the implementation of electronic technologies for catch and compliance monitoring, and improvements to fishery information systems. https://www.nfwf.org/programs/fisheries-innovation-fund
- Saltonstall Kennedy Grant Program (NOAA Fisheries). The objective of the Program is to address the needs of fishing communities in optimizing economic benefits by building and maintaining sustainable fisheries and practices, dealing with the impacts of conservation and management measures, and increasing other opportunities to keep working waterfronts viable. Recent priorities include: Directly impact fisheries through the development of science and technology based projects that lead to the strengthening of existing markets and opportunities in U.S. fishing communities. https://www.fisheries.noaa.gov/grant/saltonstall-kennedy-grant-competition
- America's SEED FUND. Run by the United States National Science Foundation, the SEED FUND helps startups navigate the earliest stages of technology translation (up to $2 million per pitch over 36 months) to create new services and scalable solutions that meet an important, unmet need. This would likely require the Data Portal to scale up to provide access to all regional fisheries or be the foundation for a country-wide data initiative. https://seedfund.nsf.gov/apply/get-started/
- Environmental Protection Agency National Environmental Information Exchange Network Grant Program. This program is designed to encourage state and other partners' data-integration efforts. Funding will be provided through grants to states, so would require expanding the portal to include New England states data and participation. https://www.epa.gov/exchangenetwork
- USGS Community for Data Integration. A dynamic community of practice working together to advance scientific data and information management and integration capabilities in the USGS. They annually fund projects that promote data integration for interdisciplinary research, innovative data management, and demonstrations of new technology. https://www.usgs.gov/centers/community-for-data-integration-cdi
- Spherical Analytics/Grantham Trust. Spherical Analytics builds trusted platforms for global environmental data and insightful analytics through big data, AI-

generated analytics, and blockchain technology. While they do not offer funding, their platform and resources could be leveraged to build out the data portal. Additionally, they may be able to provide funding mechanisms and funding sources to accomplish the project. https://www.sphericalanalytics.io/

- Kingfisher Foundation. They invest in and promote innovative ideas and practices to align economic incentives and public policies with sound fishery management and effective conservation. http://www.kingfisherfoundation.org/
- The Environmental Systems Research Institute (ESRI) Conservation Program offers sources of technical and other assistance to create and develop spatial analysis, computer mapping, and geographic information systems (GIS) capabilities among nonprofit organizations and individual projects of all sizes and types worldwide. https://www.esri.com/en-us/industries/conservation/overview
- Walton Family Foundation. Proposals are accepted by request only, but Walton's 2021-2026 domestic ocean strategy is focused on innovations in electronic monitoring and electronic reporting technology and data management, as well as building climate and socio-economic resiliency for fishing communities. A deeper integration of climate data into the portal to correlate with fishing trends could be of interest to this funder. https://www.waltonfamilyfoundation.org/

Funding the ongoing maintenance of a groundfish data portal is challenging. The fishery is barely profitable and the government continues to subsidize monitoring through Congressional appropriations, in an effort to create the accountability necessary to rebuild groundfish stocks without further gutting the dwindling fishing effort. Despite the recent inflation rates, ex-vessel prices to the fishermen have remained stagnant. Sectors managers are facing challenges in funding the sectors through landing and lease fees, and they are a regulatory requirement. Adding additional fees to the fishing businesses is a non-starter at this time.

APPENDIX 1: INDUSTRY USER CASE SCENARIOS

Introduction

Critical to planning a data portal interface is understanding the variety of end users, their preferences, and how they will use the technology. To accurately describe the variety of fishing industry end users, the program used diverse methods to engage with commercial fishermen and other industry members (sector managers, service providers) to develop Industry User Case Scenarios.

Feedback and discussion occurred with more than 26 unique industry individuals through the following mechanisms (several participated multiple times):

1. Electronic Survey to gauge data priorities and interest in data sharing
2. Three Workshops with short visual presentation of examples and in-depth discussions on portal data sets and functionality.
3. Functioning online Data Portal Prototype to test, loaded with mock data (available online for four months)
4. Electronic Survey to gauge response to Data Portal Prototype
5. Individual conversations between project staff and end users
6. In-depth pilot testing of a secure version of the Data Portal Prototype, loaded with their own real, confidential data and detailed responses

Of the 26 project participants, 20 are commercial fishermen that seasonally target groundfish, in addition to other fisheries depending on the season (there are almost no year-round groundfish-only fishermen).

Industry guidance was used to develop and refine the Data Portal Prototype and to develop six user case scenarios, detailed in the following pages.

Four of the user cases are supportive of building the portal and would use it regularly or semi-regularly, depending on how often the data is updated. One of the user cases might use it, if the design is improved to be more user-friendly and simpler. One of the user cases would not ever use the portal and is included as a reference: if this type of fisherman will never use the portal, the design does not need to accommodate their preferences.

These user case scenarios should be used to further refine the data portal during development or to support other efforts to build data/technology applications for fishermen.

Note, all users were interested in real time tracking of remaining individual quota balances; this data stream does not yet exist, except as created internally by Sector Managers. NOAA tracks remaining quota at the aggregate sector level and cannot provide a data stream for individual quota balances without overhauling their sector data infrastructure and data collection.

User Case 1: Pro Technology & Accountability

Profile

Comfortable with computers

Participates in industry research

Early adaptor of EM (audit model)

Champion of low-impact, sustainable fishing techniques

Decades of fishing experience

40-55 years old

- Would use data portal at home on Apple laptop or iPad tablet
- Priorities (in order):
    a) Quota monitoring – how much quota used per trip and remaining annual balance (will need data of record for landings and discards)
    b) Cost control: comparing eVTR and dealer
    c) Feedback mechanism for compliance with EM protocols (detailed EM data sent to NOAA plus feedback letters)
    d) CPUE + trip returns = revenue per unit effort; trends with substrate, depth, temperature
    e) Fishing planning tool: catch rates as bar charts; use environmental data find current and historical specific bottom temperatures at certain depths (+/- 5 fathoms) within 10 miles of fishing location
- Would like to customize a “quick links” menu and to change the aesthetics with personal background picture.
- Would use bimonthly if portal had near real time data (updated weekly); would use sporadically for superficial business planning if could only access historical data to compare year over year data.
- Layout is self-explanatory; intuitive options (at least on normal screen size)
- Prefers selecting specific trips to examine through the summary table/list; less data entry the better.
- Not interested in using map features.
- Would download data and save for future use and manipulation (quota records, year-end summaries, etc.)

User Case 2: Research Focused

Profile

Comfortable with tech after thorough training

Participates in Maximized Retention EM Pilot and eMOLT or Study Fleet

Uses paper logbooks and electronic chart plotter to track fishing history

Experienced fisherman

30-60 years old

- Would use data portal on home computer (PC) in conjunction with paper log books to remember past years' catches:
  - review historic trends to plan for upcoming quota leasing
  - identifying areas and times to avoid low-quota stocks
  - identify areas and times of high catch rates
- Would use regularly if portal had in-season data (updated weekly), but would use sporadically if could only access previous years' datasets.
- Would not download raw data or charts, unless needed to dispute a data issue with sector manager or government.
- Priorities:
  - Accessing own eMOLT temperature probe data in the portal, to overlay with catch data is a very high priority. Prefers to see temperature and catch data over time on the same graph (double Y axis).
  - Summarizing by market categories in Dealer Landings to track landings of sublegal fish (landed as part of EM program).
  - Price paid by dealer for each species/market category on a given trip (if not individual data, then a weekly regional average)
  - CPUE trends to identify high and low catch rates (unit= hours towed)
- Not interested in EM feedback charts – reading the feedback letter is enough.
- Not interested in VTR/dealer comparison.
- Appreciates clear instructions and video demonstration of how to use portal.
- Not interested in sharing data, with exception of accessing environmental data

User Case 3: New Captain

Profile

Comfortable with computers

Recently became owner/operator of their own fishing vessel

Actively building knowledge base of fishing trends and environmental correlations

Has latest technology in the wheelhouse

28-35 years old

- Will use data portal regularly, on home computer or, if available, on tablet or phone.
- Interested in sharing their own and accessing other fishermen's fishing data, as long as confidentiality is maintained.
- Extremely interested in overlaying substrate, depth and environmental data with catch data, to help identify trends in fishing efficiencies.
- Projections of future business success, backed by reports detailing revenue history plus landings records, would be used to help make business investment decisions, including seeking loans/capital.
- Expects to be able to track quota usage in real time with phone based application.

User Case 4: Average Fisherman

Profile

Doesn't like computers, but can use after thorough training

Uses paper logbooks and electronic chart plotter to track fishing history

Experienced fisherman

30-60 years old

- Might use data portal on home computer or Android tablet in conjunction with paper log books to remember past years' catches:
  - review historic trends to plan for upcoming quota leasing
  - identifying areas and times to avoid low-quota stocks
  - identify areas and times of high catch rates
- Would use regularly if portal had in-season data (updated weekly), but would use sporadically if could only access previous years' datasets.
- Accessing own eMOLT temperature probe data in the portal, to overlay with catch data is a very high priority. Prefers to see temperature and catch data over time on the same graph (double Y axis).
- Comfortable using Fish Online eVTR app but not FLDRS (simpler is better)
- Needs clean, clear report templates for downloading data; will not download and manipulate raw data.
- Doesn't know how to access the limited electronic data that is already available to captains online (VERS, Fish Online)
- Wants to customize with a "favorites" list on the dashboard: shortcuts to key features.

User Case 5: Old School Captain

Profile

Dislikes computers

Struggles with mandatory electronic reporting

Experienced fisherman with decades on the water

Uses paper logbooks to track catch and fishing history

50-70 years old

- Prefers paper records to electronic records – data portal provides no value to fishing business.
- Won't use data portal at all unless adds a lot of significant value, specifically:
  - tracks individual quota balance in real time, and
  - displays regional weekly ex-vessel prices compared to own ex-vessel prices.
- Believes they already know/understand how tides, weather, and temperature impacts fishing; doesn't see value in comparing own data to environmental data sets.
- No interest in sharing their own data with the fishing community, nor accessing other fishermen's data
- Doesn't know how to access the limited electronic data that is already available to captains.

User Case 6: Sector Manager

Profile

Professional staff for industry cooperative

Proficient with computers and data

Manages landing and discard data for fishermen members to report quota in aggregate to the federal government

Tracks each permit's quota allocation

- Will need to access their members' data through the portal, individually and in aggregate (treat sector manager like a fleet manager)
- Will use data portal regularly, on work computer, to look up details and reports for fishermen members.
- Already has access to the data in raw format, but has to aggregate, analyze, and summarize in excel; would prefer automated data portal application
- Interested in running data portal analyses and reports at different resolutions that adjust visual outputs to anonymize data:
  - Individual fishermen
  - Aggregated by small groups of fishermen (i.e., all hook gear vessels)
  - Aggregated by entire sector
- Will track where, when, and what is being caught, overlaid with environmental data
- Will compare eVTR and dealer landings to check for over/under reporting or errors.
- High priority is tracking the remaining quota available in the sector, both in aggregate and for each individual fishermen.

APPENDIX 2: SCIENTIST USER CASE

Introduction

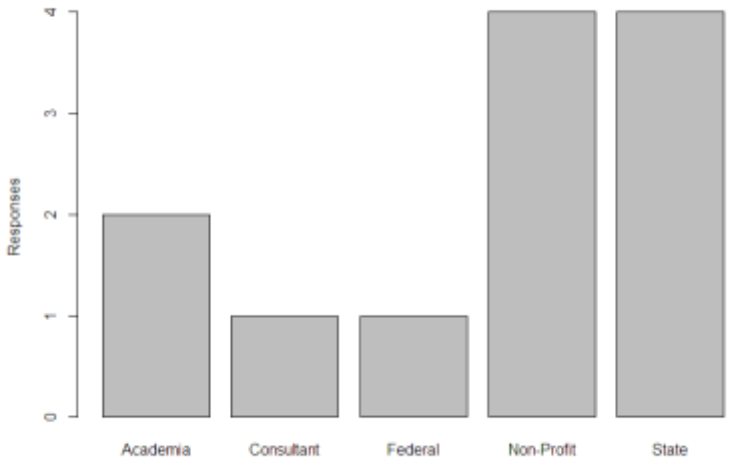
Fishermen are not the only ones interested in using the data collected on commercial fishing trips. Scientists working on fisheries analyses often rely on fishing data, if and when they can access it. Aggregated fishing data from federal and state databases is freely available at Atlantic Coastal Cooperative Statistics Program (https://www.accsp.org/), but the resolution is often not adequate for finer-scale analyses. Non-profit research organizations, academia, environmental consultants, students and government agencies have to submit data requests to NOAA Fisheries to access trip level data, which is stripped of any identifying information. This data can only be released to these non-NOAA scientists if it doesn't violate the rule of three: there needs to be at least three fishing businesses contributing to the requested data set. If landings data are involved, there also needs to be at least three dealers included in the data set.

While the rule of three is incredibly important to protect confidential business information for the fishing industry, it can make it impossible to get the resolution of data necessary for specific analyses. For ports that only sell their fish to one or two dealers, this means that multiple ports have to be aggregated until you reach three dealers. Or if trying to tease out weekly or monthly trends by gear type in a specific area, there may not be three fishermen for each gear type for each week. It can also be difficult to find lists of available data fields and understand exactly what needs to be requested in order to get the most out of the data. Some scientists work directly with fishermen partners on analyses and get high resolution, vessel specific data by having the fishermen sign data sharing agreements to access just the fishing partners' data.

It is also not uncommon for data requests in New England to take several months to complete, creating delays in research projects. Since scientists face similar hurdles to fishermen in accessing fishing data, scientists were included in the user case scenario planning for the New England Groundfish Data Portal.

Methodology

Academic, non-profit, state, and federal scientists throughout New England were surveyed on their interest in and use of fishing data, with particular focus on the novel electronic monitoring (EM) data stream. Many individuals represented larger institutions or lab groups that could make extensive use of these data. The twelve respondents come from a range of sectors including non-profit research organizations, government agencies, academia, and environmental consultants (Figure 1). They completed the multiple choice survey and provided written text responses.



Most of the interest has come from non-profits with a history of working directly with the fishing industry as well as state regulatory agencies. Presumably, scientists working at federal agencies already have access to this information, and academic scientists more often generate data for short-term projects rather than mining existing sources.

Results

All respondents were interested in accessing text data generated from Electronic Monitoring reviews, with most (75%, n=9) willing to accept data that was rolled up a rounded decimal degree and half being willing to work with data that was summarized by month. Most respondents indicated that they currently have no access to these data (67%, n=8).
Only 33% of respondents were interested in accessing dealer data, although this number jumped to 67% if the data could be tied to electronic monitoring data. Of those interested in dealer data, 57% would like to access dealer data at the vessel level and 57% could use it at the port level. About 60% of respondents were interested in Electronic Vessel Trip Report data, although this number jumped to 85% if eVTR data could be linked to electronic monitoring data. Finally, 67% of respondents were interested in temperature probe data, but 92% were interested if it could be linked to electronic monitoring data.

Detailed results are included in the Aggregated Responses section.
These responses clearly show that

1. Electronic monitoring is a valuable data source in its own right
2. Electronic monitoring data are not widely available
3. Being able to link electronic monitoring data to other data streams enhances the value of those other data streams to the scientific community.

Further discussions with scientists was limited, as the fishermen participating in the portal design were not interested in pooling their data to share among themselves or with scientists. Without their interest in sharing data, it was premature to thoroughly vet how scientists would access and use the data portal.

Recommendations

- The ability to link datasets (with unique trip identifiers) is critical for efficient analyses and for scientists to find value in novel data streams.

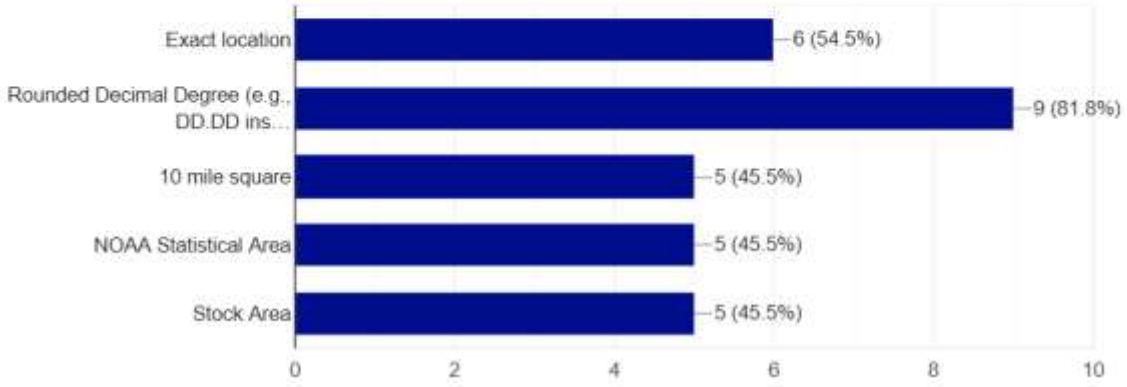
- Streamlined data requests and timely delivery of datasets should be prioritized. This could include:
  - Complete metadata available prior to request, to ensure request of correct data fields.
  - Standardized field formats across data sets.
  - Standardized online form for requests with drop downs of available data sets and data fields.
  - Automated processing of requests.
  - Guidelines for how to maximize data resolution while meeting rule of three.
- Identify opportunities to provide anonymized data with high temporal and spatial resolution (at the collection level: hours/minutes and exact location (decimal degrees), although most can work with rounded decimal degrees (DD.DD) or 10 mile squares.
- Identify solutions that provide both fishing and temperature probe data collected on the same trip.

Aggregated Responses

Electronic Monitoring Text Data (Deeper Dive)

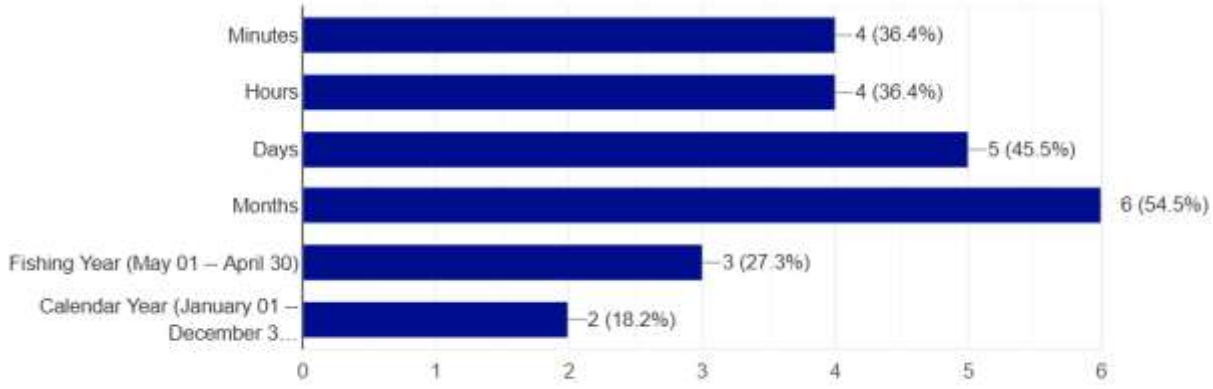
At what spatial resolution are the data useful to you? In other words, would the data only be useful if you knew the exact location where it was collected? Could it be rolled up to a broad stock area and still be useful? (check all that apply)

11 responses



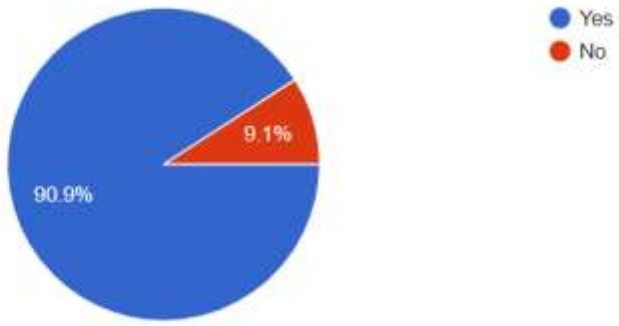
At what temporal resolution are the data useful to you? In other words, would the data only be useful at a minute by minute resolution? Could it be aggregated to a fishing year or calendar year and still be useful? (check all that apply)

11 responses



Would the data be useful if they were only available aggregated by gear type? (i.e., individual vessels were not identifiable)

11 responses



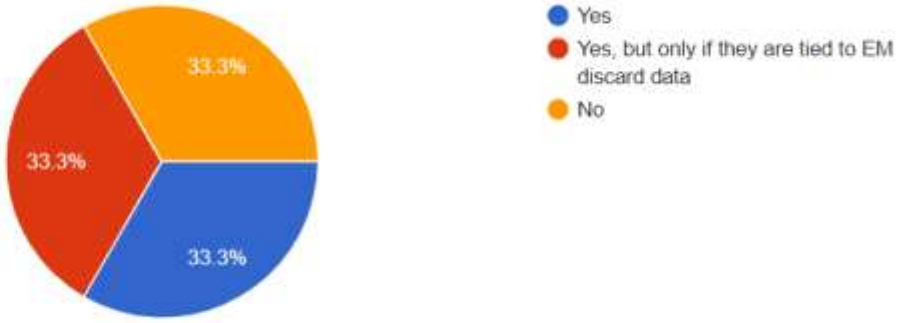
How would having access to EM data impact your research program?

- Discard estimation and fishery catch rates
- It would allow better assessment of populations and potential impacts of pending projects
- Provide my biologist colleagues (not me) with data for correlation w/environmental changes
- Will allow me to think of ways to incorporate EM data in stock assessment and better understand fleet dynamics
- We would like EM combined with logbooks on Party charter vessels to eliminate dockside interviews and FHS calls.
- Possible to evaluate effects of bycatch reduction gears/devices in use
- Not needed--we'll just do with what we've got
- Increase data available for comparison
- Help to put local (state waters) survey results into perspective with more regional data

Dealer Data

Are these data of interest to you?

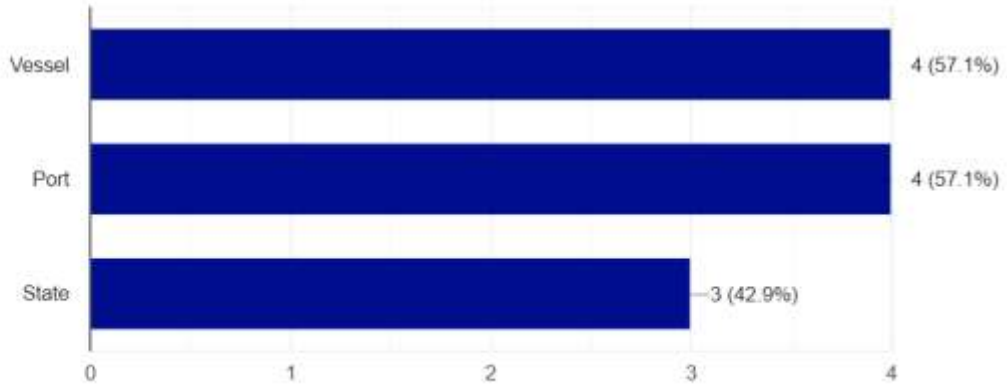
12 responses



Dealer Data (Deeper Dive)

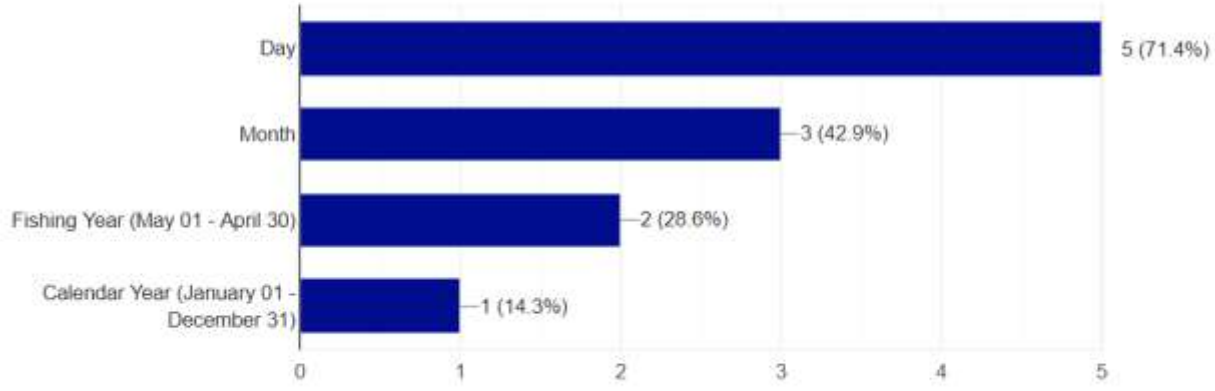
At what spatial resolution are the data useful to you? In other words, would the data only be useful if you knew that it was from a particular vessel? Could it be rolled up to a port or state and still be useful? (check all that apply)

7 responses



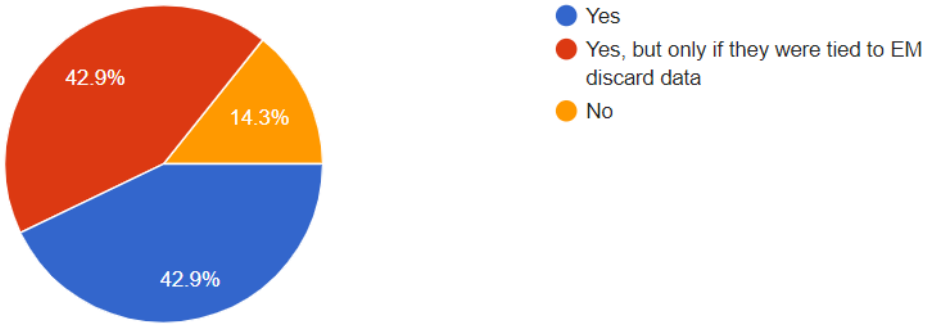
At what temporal resolution are the data useful to you? In other words, would the data only be useful at a daily resolution? Could it be aggregated to a fishing year or calendar year and still be useful? (check all that apply)

7 responses



Would the data be useful if they were only available aggregated by gear type? (i.e., individual vessels were not identifiable)

7 responses



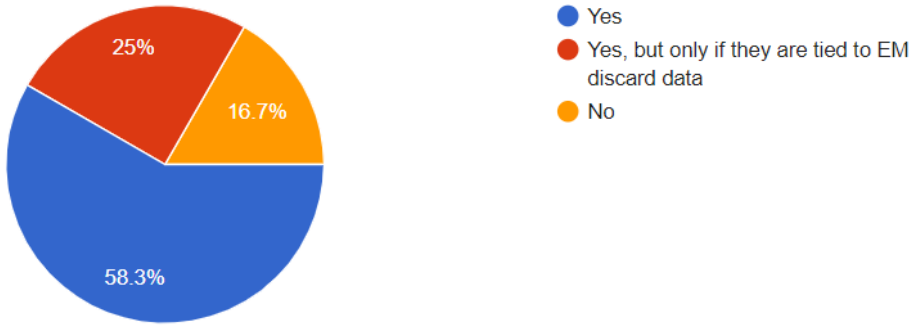
How would having access to dealer data impact your research program?

- Landings estimates and landings per unit effort
- Cross check data quality, better estimate landings and uncertainty
- Better decision making about spatial management, better identification of user conflict

Electronic Vessel Trip Reports

Are these data of interest to you?

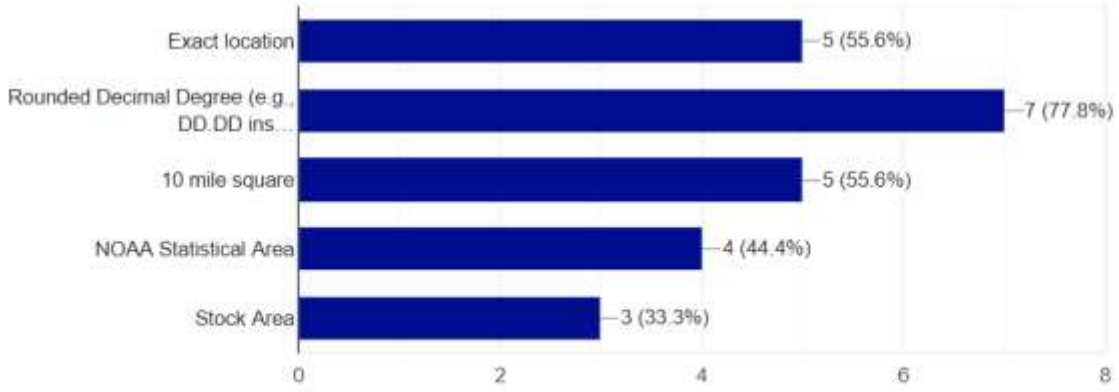
12 responses



eVTR (Deeper Dive)

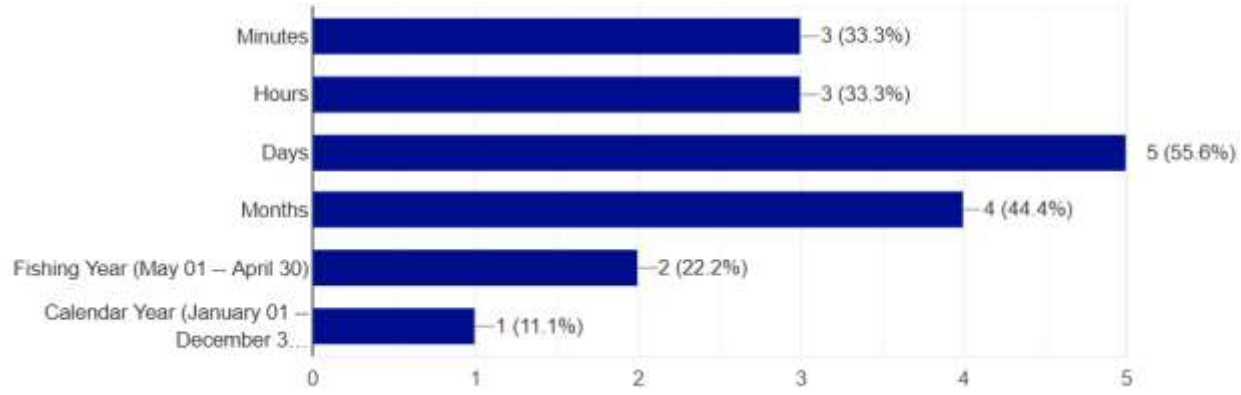
At what spatial resolution are the data useful to you? In other words, would the data only be useful if you knew the exact location where it was collected? Could it be rolled up to a broad stock area and still be useful? (check all that apply)

9 responses



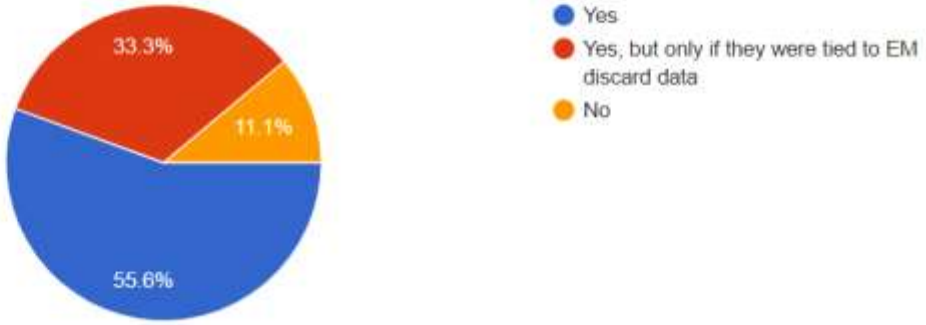
At what temporal resolution are the data useful to you? In other words, would the data only be useful at a minute by minute resolution? Could it be aggregated to a fishing year or calendar year and still be useful? (check all that apply)

9 responses



Would the data be useful if they were only available aggregated by gear type? (i.e., individual vessels were not identifiable)

9 responses



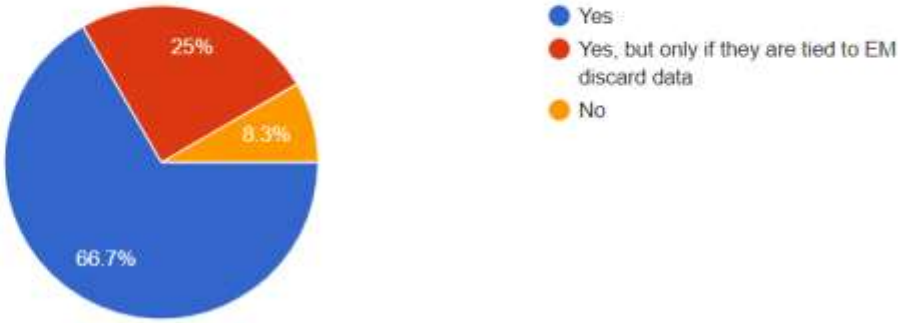
How would having access to eVTR data impact your research program?

- Estimating fishing effort, location of catch and fishery catch rates
- It would provide my biologist colleagues w/catch vs temperature correlations
- Better understand CPUE and fleet dynamics, making it possible to use the data in stock assessment
- eVTR data is critical for state management, we work with this data every week for quota management

Temperature Probe Data

Are these data of interest to you?

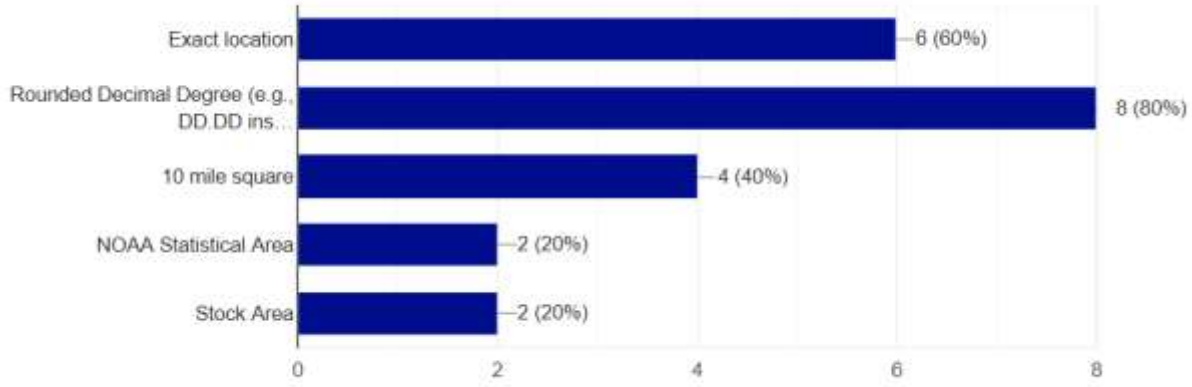
12 responses



Temperature Probe Data (Deeper Dive)

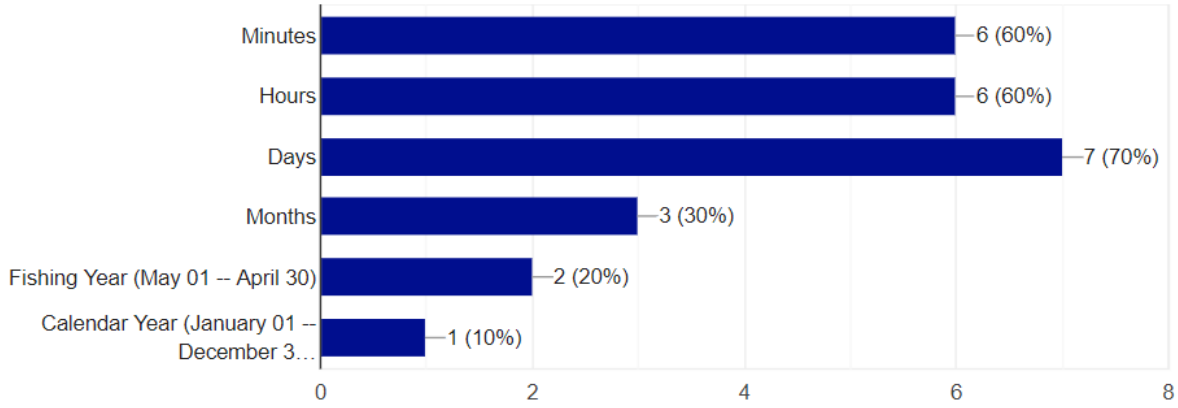
At what spatial resolution are the data useful to you? In other words, would the data only be useful if you knew the exact location where it was collected? Could it be rolled up to a broad stock area and still be useful? (check all that apply)

10 responses



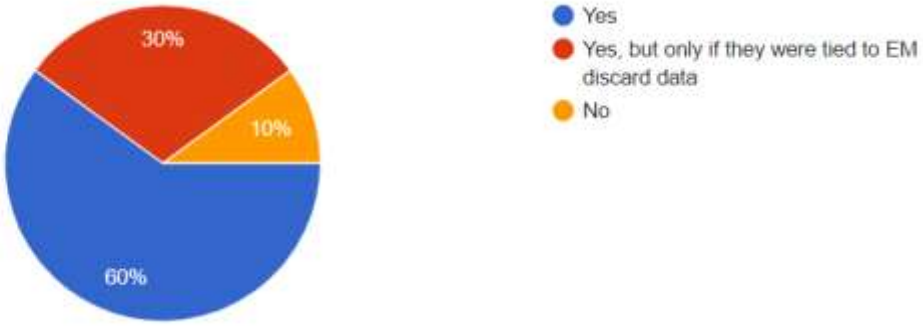
At what temporal resolution are the data useful to you? In other words, would the data only be useful at a minute by minute resolution? Could it be aggregated to a fishing year or calendar year and still be useful? (check all that apply)

10 responses



Would the data be useful if they were only available aggregated by gear type? (i.e., individual vessels were not identifiable)

10 responses



How would having access to temperature probe data impact your research program?

- Validating fishing effort events, environmental effects
- Combining eMOLT data with the many other bottom temp observations is our goal
- Better understand thermal habitat and impacts of climate changes
- Better habitat mapping
- Increased spatial scale and sample size to look at trends we are examining in our data
- Modeling habitat requirements for species requires temperature data matched to catch at a minimum. Very important for future fisheries considerations

Is there anything more you would like us to consider that we didn't ask?

- I strongly support this initiative! Disaggregated data are much more informative, but aggregated data are still useful"
- How do we fund more temperature probes on more vessels?

- NOAA Fisheries is adding too many gear metrics to the eVTR. It will ruin the initiative. Only the basic gear data should be collected on a eVTR for a trip (hauls, amount of gear fishing) and the additional gear info (Rings size, mesh size, #panels, pots per string, mesh size etc) should all be collected via a survey with a frequency by wave (2 months). The extra gear metrics have crashed the system several times and have resulted in fishermen becoming increasingly frustrated and disenfranchised with the eVTR system. KEEP IT SIMPLE STUPID!
- Gear descriptions may need to be more detailed than currently reported. Codes are somewhat limited, but might be useful if they are for bycatch reduction (e.g. Ruhle trawls)
- Would this ever be a potential source of tissue samples? Fin clips, etc.
- What about scientifically collect biological data from commercial fishing trips (ASM, observer data)? Is this about groundfish "fishery dependent data that is available in New England" or just about electronic monitoring/fishermen reported data?