Global Maritime Forum

REPORT ON THE WORKSHOP

National Maritime Intelligence-Integration Office
2017 Global Maritime Forum Report

The Evolution of Big Data and Opportunities for the Maritime Domain:
The Importance of Public-Private-Academic Partnerships

Held at
California State University Maritime Academy
Vallejo, California
13-14 November 2017

National Maritime Intelligence–Integration Office
Global Maritime Forum Report 2017

The Evolution of Big Data and Opportunities for the Maritime Domain: The Importance of Public-Private-Academic Partnerships

Event location: California State University Maritime Academy, Vallejo, CA
13 & 14 November 2017

Table of Contents

Executive Summary

The National Maritime Intelligence-Integration Office (NMIO) hosted its sixth annual Global Maritime Forum (GMF) from 13-14 November 2017, at California State University Maritime Academy (Cal Maritime) in Vallejo, CA. The GMF supports and aligns with the priorities established by the Director of National Intelligence for NMIO, the National Intelligence Manager for Maritime.

NMIO Priorities

- Develop the Global Maritime Community of Interest (GMCOI)
- Improve information/intelligence sharing
- Advocate GMCOI collection and analytic priorities
- Integrate Science and Technology

The theme of the 2017 forum was the "Evolution of Big Data and Opportunities for the Maritime Domain: The Importance of Public-Private-Academic-International Partnerships." The workshop was a series of plenary panels and collaborative break-out sessions focused on the early stages of enhancing maritime security by improving and developing big data collection, availability, and access for all maritime stakeholders.

Attendees from across the GMCOI included over 130 stakeholders from United States Government (USG), private sector, academia, Non-Governmental Organizations (NGOs), and international subject matter experts. Participants from outside the United States included Canada, Chile, Germany, Italy, Japan, the Philippines, and the United Kingdom.

A main objective of the forum was to enable development of a maritime common operating picture (COP) fed and built from numerous data sources. The COP could take many forms, for example, a live digital map automatically updated in real time. Through increased situational awareness, maritime security is significantly enhanced when all stakeholders are provided with timely, accurate information and are thus enabled to make better, more informed decisions. The subjects of such data include timely updates to navigation hazards, criminal activity, piracy, terrorism, and vessels in distress. The premise that each stakeholder is developing, has access to, or is collecting unique information and including that information, adds to the granularity of a COP, thereby making it more valuable for all stakeholders. Each additional piece of information would in theory make the COP more accurate, and would enable other stakeholders to make better informed decisions in the context of maritime security.

A significant takeaway for all attendees was that a great amount of mature and fielded technology already exists for collecting maritime information. The main challenge in light of that realization is that such information is not being effectively aggregated, shared, analyzed, or understood as efficiently as possible. Significant portions of information being collected by individual stakeholders cannot be

accessed by other stakeholders due to policy restrictions of a particular stakeholder, as well as data formatting differences. Policy restrictions and data formatting differences typically exist because the information was collected by systems developed independent of one other or may contain intellectual property that is private, proprietary, or sensitive. Even if the challenges of policy and data formatting differences were overcome, the lack of common software to read the data remains an issue.

The need for advancing partnerships that enable collaboration, sharing, availability, and accessibility of data was widely discussed and agreed upon.

In order to advance these partnerships, attendees shared and discussed best practices and challenges associated with big data applications and architectures. Additional challenges identified include: 1) Data management; 2) Development of algorithms for automatic aggregation and analysis of data from different sensors and formats; 3) Dissemination of timely, accurate, predictive, and relevant information; and 4) Navigating the optimal path for maximum data throughput, availability, and accessibility, while simultaneously staying inside established legal frameworks.

The GMF resulted in the successful sharing of technical information among a diverse set of governmental, commercial, and academic stakeholders as well as actionable recommendations for future efforts to enhance Maritime Domain Awareness.

GMF 2017 Success Metrics

• **Facilitate and enable operations by coordinating input from operators to influence S&T**
  ○ **Day One: Each of four panels identify cutting-edge best practices**
  ○ **Day Two: Working Groups**
  ▪ **Identify at least one capability solution for each Day One panel**
  ▪ **Develop tangible outcomes that can be used to inform NMIO's FY 2019 big data advanced analytics initiative**

GMF 2017 Overview
WHY:
FY 2019 Big Data Advanced Analytics Initiative: • **Unclassified Data & Analytics as an Enterprise Asset** • **USG-Private-Academic-International Partnership**
HOW:
2017 GMF Focus: Accessing and Managing Data **2018 GMF Focus: Advanced Analytics**
WHAT:
Day One – Accessing & Managing Data – Best Practices 1. **Managing Maritime Big Data Volume & Velocity** 2. **Aggregating Multi-Source Data in the Maritime Domain (Best Practices)** 3. **Policies Impacting Open Source and Publicly Available Information** 4. **Developing a Maritime Data Layer Cloud/Access Backbone** **Day Two – Brainstorm Solutions: What Right Looks Like**

Welcome Address

RADM Robert Sharp, Director, National Maritime Intelligence-Integration Office and Commander, Office of Naval Intelligence, opened the GMF by stressing the importance of collaboration and information sharing.

He indicated that we will never solve maritime problems on our own, and he highlighted the importance of USG working with industry, academia, NGOs, and international partners to help mitigate information gaps.

He discussed the integration role of NMIO, highlighting the importance of making information more readily available to address maritime challenges, the aforementioned gaps, and security concerns.

RADM Sharp emphasized how Big Data processing is critical to filling these gaps because it allows maritime stakeholders of all levels the opportunity to make better informed and more timely decisions. The science of processing Big Data, however, is still in its infancy—the goal now is to effectively and rapidly analyze and disseminate (or make accessible) the relevant information to the appropriate stakeholders.

Achieving these goals involves answering several key questions.

- Do we have enough data?
- How do we make sense of the data we already have?
- How can we work together to better share information and intelligence?

He also emphasized the importance of public sources of information. Identifying new methodologies or linking existing methods for sharing existing public data is a key step towards developing a COP or commonly accessible data base for GMCOI living in the unclassified space. He stressed that the GMCOI must continue to prioritize improvement and asked participants to explore how the maritime community can work together to create an unclassified data enterprise.

Day One

Keynote Speaker: Mr. Snehal Antani

Lessons Learned from Application Data Sharing Related to Disaster Relief

Mr. Antani, board member of NeverAgain.org and an executive at Splunk, shared lessons learned from his work developing an application (app) for IT and personal electronic devices to assist in disaster relief. The app provides personalized information to both victims and first responders, relying on crowd-sourced data that is aggregated and synthesized for an optimized response.

Mr. Antani provided lessons from this experience applicable to organizations looking to become data-driven. Those lessons include:

- **Poor exploitation of data:** From Mr. Antani's experience, a "Data-up" approach (that is, an approach where data is curated, stories are built from the data to gain insight, and finally a decision is made based on that insight) is the wrong approach to exploiting data. He argued for an opposite, "Decision-down" approach where the first step is to determine the decision to be made and identifying information required for said decision. Data analysis can only occur after these steps.
- **Data needs to tell a story:** Data insight should be traceable to provide confidence to the user of the information. Delivering this capability requires the right project team; such a team should include a domain expert, a data architect, data scientist, and a data journalist.
- **Data access needs to be decentralized:** There is a tendency within organizations for individuals to create empires, leading to data being centrally controlled. While data itself should be centralized, Mr. Antani stated that data access should be decentralized. Organizations should strive to break down internal cultural barriers to de-silo data and make it potentially shareable (within certain context).

Takeaways

- The next leading capability needed in data science is real-time stitching of multiple sources of data—including images, video, audio, and geospatial data for a richer information product.
- Another important capability is what he called the "orchestration of things," or being able to dynamically coordinate algorithm workflows.
- The data science community needs to focus on the data landscape and find a way to move beyond structured and unstructured data to connect the dots.

Panel One: Managing Maritime Big Data Volume & Velocity

While introducing the panel, Dr. Marcus Stefano of the Naval Postgraduate School spoke to the issues of data sensor coverage, persistence, and how the focus on the behavior of particular actors drives the need for specific data.

He discussed a lack of technical solutions that "trigger" action. His questions for the panel included:

- How do we store lots of data?
- How do we control access (both broadening and limiting, depending on data)?
- How do we "normalize" data so it can be used effectively?
- How do we move quickly between analytic capabilities? Conversely, how do you chain them?
- How do we derive sense-making from data?

Commercial Capabilities

The first panel speaker, Victor Leonard, Senior Fellow at DigitalGlobe, discussed the difference between problem solving with increased amounts of useful data versus increased data that is not relevant. Too much irrelevant data is counterproductive. In his opinion, commercial capabilities are getting good enough, fast enough, and secure enough to help address many maritime operational challenges. He also advocated using less expensive commercial collection platforms (with lower resolution) to tip more expensive platforms (with higher resolution).

The panel's second speaker, Mr. Chris Bollinger of Orbital Insight, described how software is key to making sense of data. Ingesting, parsing, applying algorithms, analyzing, and presenting data in usable formats is a vitally important process. All the commercial capabilities discussed are rapidly improving, and each contributes to solving maritime operational challenges.

Technology Providing Actionable Information

Mr. Kyle Brazil of Spire, the panel's third speaker, described how his organization extends surveillance by correlating synthetic aperture radar images to support search and rescue efforts. This assists with pinpointing vessels in distress, and also aids in identifying illegal fishing and promoting a better understanding of maritime piracy.


Dr. Ruben Sorenson

The panel's final speaker, Dr. Ruben Sorenson of Planet, described how his organization uses open-ocean monitoring to combat illicit activities by processing seven terabytes of information every day via a fully automated processing-exploitation-dissemination (PED) strategy.

Panel Two: Aggregating Multi-Source Data in the Maritime Domain

Moderated by Mr. Josh Reiter of Office of Naval Intelligence, this panel included speakers from both the United States and international partners.

The first speaker for this panel, Mr. Kurt Salchert, a retired Captain in the Royal Canadian Navy now with Beyond the Border Consulting, provided an overview of structured and unstructured data analysis and the type of data that needs to be integrated to support operations.

Decomposing Essential Elements of Information

Mr. Salchert discussed the need to decompose essential elements of information and ask relevant questions, including:

- Is the data being shared with all the people who need to see it?
- Have all the necessary connections between data been made?
- Does this represent all the data?
- What are the third or fourth order effects to the data policy?

He argued that the more transparent we are about what is being done with the data, the better the results of any associated analysis. Exploitation and processing methodologies should be repeatable with clearly identified steps. If the resulting analysis is not correct, it allows one to go back and examine the process to find the error in the analysis.

Data Desired to Support Operations

The panel's second speaker, Mr. Tony Long of Global Fishing Watch, described what data was needed to support operations to defend critical infrastructure, defeat illicit trafficking, and enhance environmental disaster impact statements.

He argued the need to develop a global list of fishing vessels and increase the licensing fee to fund enforcement.

Dr. Jill Brandenberger

Challenges to Data Fusion and Aggregation

The third speaker, Dr. Jill Brandenberger, Manager of Environmental Security at Pacific Northwest National Laboratory, noted that the challenge is not the availability of imagery but the processing, exploitation, and dissemination of that imagery.

Dr. Brandenberger further identified the need to include social media analysis to validate algorithms. She added that data ontology and error propagation remain a problem and discussed the importance of non-traditional data and what those sources of data bring to our understanding of the problem.

LCdr Michael Bielby of the Government of Canada was the panel's final speaker, and he discussed the importance of being able to aggregate data to support data analysis and visualization and to make it available to partners. In this context, he argued the need to collect everything.

Panel Three: Policies Impacting Open Source and Publicly Available Information

Building Awareness Through Exploitation of Open Source Information

This panel was moderated by Dr. Paul Shapiro of the National Defense University, who briefly discussed the big data-related disruption to maritime transportation. He argued the need to develop the maritime equivalent of the human genome project.

Dr. Assis Malaquias of Cal Maritime described information sharing taking place in Africa. He went on to say that maritime security requires multilateral coordination to address a variety of common threats and challenges, which includes illegal, unregulated, and unreported (IUU) fishing. He indicated that the likely reasons contributing to maritime security challenges include:

- Poor governance
- Little or no maritime domain awareness
- Lack of capacity to effectively extend sovereignty to the maritime domain

The Djibouti Code of Conduct and the Yaoundé Code of Conduct are two examples of current working solutions in the Indian Ocean and Gulf of Guinea, respectively.

Mr. Brendan McCahill

Mr. Brendan McCahill of Descartes Systems Group described how his organization is able to track imports and exports of key products from the United States, Latin America, Asia, and the European Union to monitor commercial competitors in strategic markets. Descartes has proprietary information allowing them to screen supply chain actors and customers against global restricted-party information, thereby minimizing bribery and corruption risk. Mr. McCahill suggested that the maritime community should continue to mine raw data. Raw data can be valuable as a stand-alone product or it can provide added value when combined with other data streams.

Ms. Margaret Smith of Accenture Federal Services described how her firm exploits open source and publicly available information to detect trafficking activity by using the combination of a fraud classification algorithm, location inference, and social network analysis.

Mr. Patrick O'Keeffe, NATO Legal Advisor at the Centre of Excellence for Operations in Confined and Shallow Waters, was the panel's final speaker, taking GMF participants through the global challenges of a cross-domain environment as it relates to sovereignty in cyberspace.

Panel Four: Developing a Maritime Data Layer Cloud/Access Backbone

Processing Exploitation and Dissemination

This panel was moderated by Dr. John Mittleman of Naval Research Laboratory.

Julie Baker, Vice President of Operations and Co-founder of Ursa Space Systems Inc., argued that having the right sense-making tools is key to building a successful infrastructure for big data analytics. Building a successful infrastructure is important because there will never be enough analysts to analyze all collected data. Therefore, Ms. Baker urged the maritime community to find ways to automate as much of the PED cycle as possible.

Dr. Pavel Machalek, Co-founder and Chief Executive Officer of SpaceKnow Inc., described how his firm uses deep neural networks—an artificial neural network used in machine learning to model complex non-linear relationships—to capture hidden underlying patterns in provided training data to predict future activity using previously unseen data.

Structured Data

Dr. Andrew Gearhart, Senior Data Scientist at The Johns Hopkins University Applied Physics Laboratory, described the limitations resulting from organizations not having the interest, time, or capability to fully discover the many artifacts that can exist in a single data source. Although he was generally skeptical of ontologies, he argued the need for a meta-ontology to relate different data repositories to one another.

The panel's final speaker, Pieter Decker of Harris Corporation, described our tendency to build systems first and then figure out what to do with the resultant information. He noted that 95 percent of all data is unstructured and that this data must first be structured to be usable.

Lunch Keynote Address: Dr. Gunnar Carlsson

Complex Data and Artificial Intelligence

Dr. Carlsson, President and Co-Founder of Ayasdi and Professor Emeritus at Stanford University, spoke about the importance of learning from complex data. He argued that the complexity of the data, not the amount of data, is a fundamental issue in big data analytics. He added people want more than mere insight. Information should inform a decision and allow an action to be taken.

Artificial intelligence may allow development of actionable intelligence to address real-world problems. Dr. Carlsson distinguished between narrow AI, where human tasks are automated, and broad AI, which follows the path of discovery, prediction, justification, action, and finally learning.

The challenges with the "justification" step can be partially overcome by looking at data shape distinct from data visualization. In this context, the described topological data analysis, using unbiased, unsupervised, discovery algorithms can be used to identify clusters of similar data points. This results in a justifiable visual presentation showing the discovered relationships in the data, with built-in resolution rheostats. In a given scenario, this allows one to examine the data with either more or less fidelity as required.

Dr. Carlsson concluded with the need to be more flexible when modeling real-world problems and not focus on the "gee whiz" problems, as they are not always the most important.

Day Two

Dr. Lilian Alessa

Cognitive Capacity for Accurately Processing and Interpreting Information

Dr. Alessa, the Deputy Chief, Global Strategy, Policy, and Plans at the Department of Homeland Security and President's Professor at the Center for Resilient Communities, University of Idaho, discussed the importance of being able to answer the problems presented by data volume, velocity, and variety (complexity). Making all available data discoverable and accessible for analysis helps. However, data sharing among Private-Public Partnerships requires innovative agreements throughout the federal, state, local, tribal, territorial, and private partners when analyzing data, structured data is often limiting. A better solution is to develop algorithms able to accommodate multiple data formats.

Nonetheless, the data analyzed must be relevant and applicable for tactical decision support. The complexity of human networks and the operational environment give rise to unexpected outcomes often missed by standard linear algorithms. Further, the velocity of changing circumstances often results in a contextual drift of the operational environment which occurs faster than can be represented in data repositories.

For this reason, Dr. Alessa noted we must keep in mind the limitations of cognitive capacity over big data. The human mind only processes a few key indicators at a time. Data is scalar, having multiple characteristics of time, location, and motion that are often missed by human cognition. Decisions are often made on perceptions of information vice deep knowledge of the data. Cognitive saturation can cause technology induced environmental dissonance. This dissonance leads to poor decisions about the relevance of situational awareness in tactical context. Saturation can also result in cognitive dependency on technology. This dependency factor leads to disregarding new time-sensitive information that isn't processed within the technology, and also results in sub-standard decision-making in tactical situations. It is for this reason that not all data visualizations are useful as tactical decision aids.

Of primary note in Dr. Alessa's presentation was the need to acquire local context for higher-fidelity interpretation of data. Relying on social media for immediate context can also be misleading. The medium

only represents a subset of a population—one with a particular mind-set about being digitally connected. In the end, context is always critical to the understanding of data, and the best context is always gained from data provided by trained local observers. Unfortunately, the USG often does not structure data collection for variety or diversity in order to be able to track key indicators, lacking an overall framework for collecting observation data.

Takeaways

- Access to data can be complicated by policies that do not follow allowances in law
- Volume of data can be complicated by low grade algorithmic processors
- Variety of data can be complicated with poor quality data, such as social media
- Velocity of data in real-time context can lead to poor tactical decision making
- With few exceptions, tactical decisions must be informed by real-time context from expert local observers

NMIO Data Challenge – Accurate Assessment Of Illegal, Unreported And Unregulated Fishing

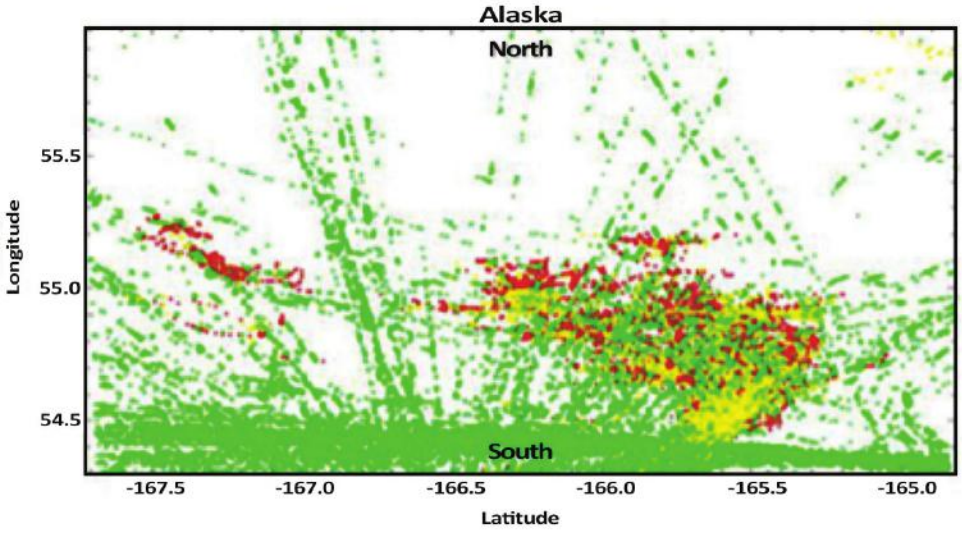
Dr. John Mittleman briefed the results of the 2015 NMIO Data Challenge. The challenge invited individuals to develop an algorithm to identify illegal, unreported and unregulated fishing. The Data Challenge was developed from the 2015 GMF, held at National Aeronautics and Space Administration (NASA) Ames in California. Answering the challenge, 119 participants from 33 countries responded.

The premise was to combat IUU fishing by identifying those who: 1) circumvent conservation and management measures; 2) avoid costs associated with sustainable fishing practices; and 3) derive economic benefit by exceeding harvesting limits. NMIO partnered with NASA and other government, industry, and academic partners to create this two-phased challenge.

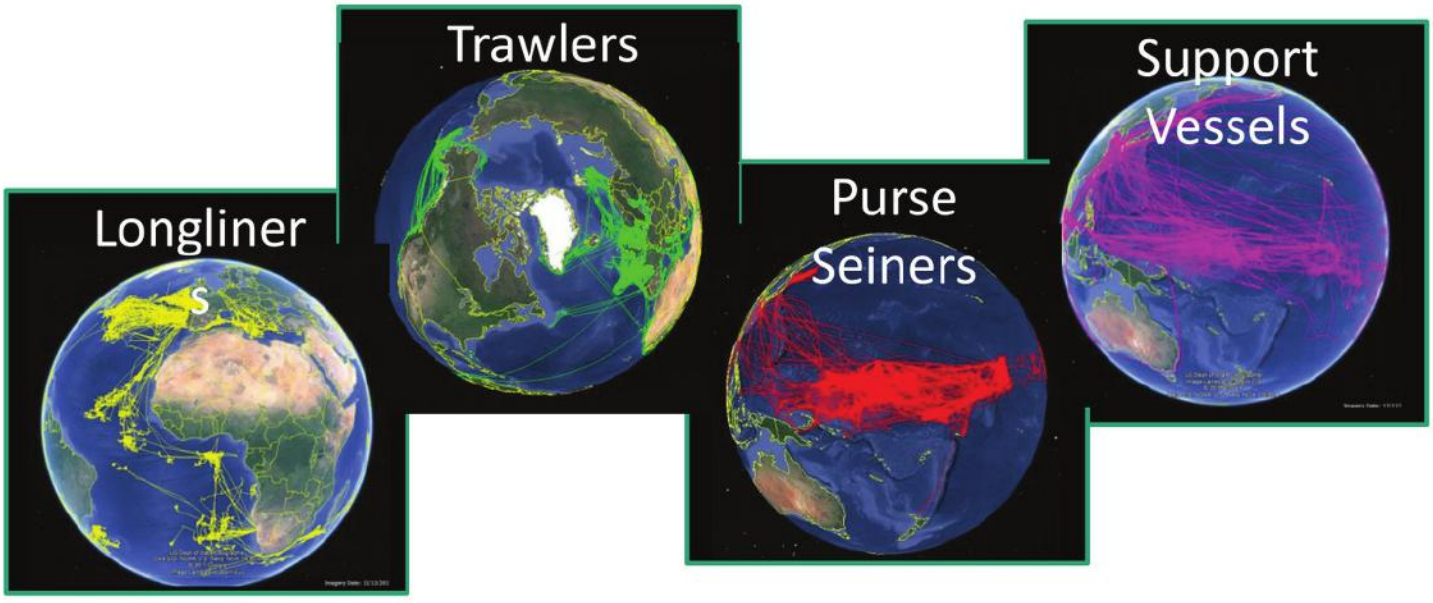
Phase One asked participants to determine, using Automated Identification System positional data, whether a vessel was fishing or not fishing. The determination would (or would not) be corroborated by maritime analysts separately establishing ground truth. This resulted in a winning algorithm that successfully processed the test data with at least 98 percent accuracy.

Phase Two participants were asked to determine, given that the vessels were believed to be fishing, what type of fishing the vessel was actually engaging in. This is important to determine if a declared fishing vessel is engaging in unauthorized fishing activity (e.g., a vessel that is declared to be engaged in long-line fishing is actually performing a type of fishing that results in significantly increased yields, and thus contrary to sustainable fishing practices, causing harm to the maritime environment).

A participant from Poland, the winner of Phase One, created an algorithm which proved over 98 percent accurate. The Phase Two winner, from Brazil, produced an algorithm that proved between 92-98 percent effective, depending on the type of fishing activity it sought to identify. Both results exceeded the competition's expectations. The winning algorithms and training data were made available to the public.



Winner, Phase One: Vessels Found Fishing vs. Those Found Not Fishing.



Winner, Phase Two: Type of Fishing Vessel

The NMIO Data Challenge encouraged a public-private-commercial partnership to expand analytic capabilities and capacity. Tapping into a broad and deep international talent pool, with no restrictions on participation, is a model to be used for future challenges.

NMIO Big Data Initiative Briefing – Fiscal Year 2019

NMIO is currently working on a Fiscal Year 2019 Big Data Initiative intended to create an unclassified data lake and make data available as an enterprise asset. Mr. Todd Boone, Head of NMIO's Intelligence Integration Department, briefed the results of a cognitive neural network prototype for multiple intelligence (multi-INT) maritime data.

Key takeaways:

- **Architecture:** Calculating gigaflops per watt, per compute-node, is a discriminator for efficient hardware.
- **Software:** Industry has developed software to implement heterogeneous computing and compute-node vectorization, to further enhance compute-node efficiency and is cheaper than changing or replacing hardware. Leading industry software-based hardware virtualization leads to reduced dependencies on architecture, and frees up the ability to use different analytic engines, which is normally tied to the type of hardware available. Software is now available to frame a fully services-oriented architecture agnostic of hardware, data, analytics, and visualization services.
- **Data:** To achieve greater insight than currently available, data should be viewed as a national asset vice an agency or office-centric asset. According to appropriate attribute based access controls, it should be discoverable, accessible, and analyzable across the enterprise. Again, industry has developed software available to make this capability a reality. Enterprise-level data purchasing and sharing agreements need to become the normal modus of business.
- **Prototype:** In a mere few days, NMIO's contractors generated a modular demonstration prototype using a cognitive neural network to analyze multi-INT maritime data without fusing the data or creating a new database from source data. Streaming meteorological data was analyzed in real-time eight different ways with streaming AIS data across sixteen different vessel and meteorological characteristics.
- **Results:** For every vessel in the AIS data feed, the results produced a safety of life-at-sea risk assessment in real-time. This prototype proved that neural networks can generate analytic results by comparing different types of data—without moving the data, keeping the data, or creating a new database from the data.

Collaborative Breakout Sessions


Ms. Mekisha Marshall

Participants were divided into small groups and asked to develop an achievable plan to influence maritime technology innovation in an identified area of concern while leveraging public-private-academic-NGO-international partnerships. Groups were asked to:

- Consider big data and velocity.
- Identify best practices for aggregating multi-source data.
- Identify policies impacting open source and publicly available information.
- Isolate how to develop a maritime data layer cloud/access backbone.

Group Recommendations and Ideas

Group One:

Hold a Phase Three of NMIO's Data Challenge focused on distinguishing fishing activities as legal vs. illegal to include international partner NGO and industry data.

Group Two:

Enable vetted access by GMCOI members to real-time data from multiple sources: government, industry, etc. Hold a foreign port data-labeling challenge.

Group Three:

Model after the Dark Web, a global, multi-user data access structure to identify tipping points of data acquired and identify when there is a quorum with sufficient data. A key question would be how to bring together all relevant maritime data.

Group Four:

Hold a big data challenge focused on illicit cargo and interdiction to get a handle on diversity of vessel types and illicit behavior.

Group Five:

Develop a single cloud-based database management system, where value-added service providers can access and provide value-added services for payment or for free, to address illegal fishing problems—a virtual fusion center.

Group Six:

Develop a network of actors, identifying also information and logistics nodes, involved in illegal fishing, similar to network analysis performed for counterdrug and counterterrorism. Focus data aggregation and analysis around "use cases" to inform decision-making for specific operational questions. (i.e. How can we tip and cue surveillance before a vessel leaves port?)

Group Seven:

Define a common standard for data aggregation. Specifically, establish trusted (like Lloyd's for insurance) global entity, privacy, and sharing guidelines.

Group Eight:

Identify policy actions and priorities, and increase transparency and access to data in the South China Sea. Move the burden of care in the maritime security realm from protecting information to sharing information.

Group Nine:

Define models of trust regarding data, to include: verification of sources, quality of management and curation, level of control, and type of user's agreement.

Memorable Exercise Outcomes

These group exercises helped identify and highlight gaps to be explored for future collaborative efforts.

- A parametric database of databases which identifies datasets and key data fields relative to operational priority, accessibility, granularity, currency, etc.
- Network analysis of the bad actors and business models that make illegal fishing more profitable than legal fishing to include indications and warning of trends which could push fishermen to illicit activity due to overfishing and reduced income.
- Understand and anticipate patterns for illicit behavior at sea and in port relative to legal activity links between ship's chandlers, stevedores, ports, vessels, operators, vessel owners, and customers.
- Viability of actionable data from non-traditional maritime sources, for example, cellular phone traffic to understand ship convergence and who is where.

Moving forward, a notional framework and roadmap would include a memorandum or an agreement similar to a cooperative research and development agreement. Elements to address/determine include:

- Public-private partnerships
- Division of labor
- Collection of resources
- Collaborative action with a timeline, milestones, and budget
- Roles and responsibility
- Measures of performance and measures of evaluation/success
- Stakeholder roles, responsibilities, authorities, resources, priorities, and policies
- Additional information including chain of custody confidence, datasets, gap identification, and problem statement development
- Cost of acquisition of data base infrastructure development

The desired end-state would be a map of key ports and locations, actors, business models, and behavior patterns both in port and at sea. Goals include: 1) Indications and warning and predictive analysis which can tip and queue maritime surveillance and interdiction before a vessel leaves port; 2) Pattern analysis able to distinguish legal fishing from illicit behavior disguised as fishing; 3) Data mapping for database purposes, key data fields, importance, and relevance to maritime security categories (such as IUU fishing). This data mapping would include operationally relevant questions; 4) Database access would include user permission rights, read/write authority, level of authority.

Additional Considerations

- Sources and quality of information being ingested
- Update frequency and currency
- Volume of data (measured in bytes or records), variety, granularity, spatiality (global/regional/national/sub-national)
- Information classification (open, government owned, restricted, privately owned, law enforcement sensitive, and classified)
- Transition results from collaboration and milestone events to practice in maritime security
- How to assess results, examine lessons learned, and iterate in spiral development

Closing Remarks From RADM Sharp

The Forum closed with RADM Sharp thanking participants for coming together to discuss big data, opportunities for the maritime domain, and the importance of public-private-academic-NGO-international partnerships.

Information gained from the workshop is only as valuable as the community makes it—what is most important is what happens next. He encouraged the community to continue discussions and dialogue, bringing forth new ideas for future Global Maritime Forum discussion.

APPENDIX A: AGENDA

Date	Time	Topic
13 NOV 17	0730-0800	**Registration**
	0800–0810	**Welcome Remarks/Administrative Notes** - Ms. Mekisha Marshall, Science & Technology Advisor, National Maritime Intelligence-Integration Office
	0810-0820	**Opening Remarks** - Thomas Cropper, President of CAL Maritime, Rear Admiral, USN (Ret.)
	0820-0850	**Welcome Address** – Rear Admiral Robert Sharp, Director, National Maritime Intelligence-Integration Office
	0900-1000	**Keynote** - Mr. Snehal Antani, NeverAgain.org
	1000-1015	**Break**
	1015-1145	**Panel Session I: Managing Maritime Big Data Volume & Velocity** – Moderator - Dr. Marcus Stefano; Panelists: Victor Leonard, Kyle Brazil, Dr. Reuben Sorensen, Chris Bollinger
	1145-1245	**Lunch Keynote** - Dr. Gunnar Carlsson, President, Ayasdi
	1245-1415	**Panel Session II: Aggregating Multi-Source Data in the Maritime Domain (Best Practices)** - Moderator - Josh Reiter; Panelists: Kurt Salchert, Tony Long, LtCdr Michael Bielby, Dr. Jill Brandenberger
	1415-1430	**Break**
	1430-1545	**Panel Session III: Policies Impacting Open Source and Publicly Available Information** - Moderator - Dr. Paul Shapiro; Panelists: Brendan McCahill, Margaret Smith, Patrick O'Keeffe, Dr. Assis Malaquias
	1545-1700	**Panel Session IV: Developing a Maritime Data Layer Cloud/ Access Backbone** - Moderator - Dr. John Mittleman; Panelists: Julie Baker, Dr. Pavel Machalek, Pieter Decker, Dr. Andrew Gearhart
	1700-1715	**Wrap up and Adjournment**
	1800-2000	**No Host Social** - @ Junction Brewery and Grill

Date	Time	Topic
14 NOV 17	0800-0815	**Recap and Administrative Notes**
	0815-0915	**Day Two Speaker** - Dr. Lilian Alessa, President's Professor at the Center for Resilient Communities, University of Idaho
	0915-0945	**NMIO Data Challenge Results** - Dr. John Mittleman, Naval Research Laboratory
	0945-1000	**Assembly of Working Groups** - Break
	1000-1130	**Working Group Break Outs**
	1130-1230	**Lunch**
	1230-1315	**Cognitive Neural Network Application for Big Data, Analytic Solutions to Big Data Briefing** - Mr. Todd Boone, Head, NMIO Intelligence Integration Department
	1315-1415	**Working Group's Continue and Prepare** Out-briefs
	1415-1430	**Break**
	1430-1630	**Presentations of Working Group Findings** - Each group - 10 minutes
	1630	**GMF Wrap up and Due-Outs/Closing Remarks** - Rear Admiral Sharp & Ms. Mekisha Marshall

2017 Global Maritime Forum Group Photo