Alaska Ocean Observing System (AOOS) Data Management Review Report

COMPILED BY PETER L. PULSIFER AND REVIEW COMMITTEE

DECEMBER 2017

PRESENTED TO: MOLLY MCCAMMON AND THE AOOS BOARD OF DIRECTORS

http://www.aoos.org
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>TABLE OF CONTENTS</td>
<td>1</td>
</tr>
<tr>
<td>PARTICIPANTS</td>
<td>3</td>
</tr>
<tr>
<td>REVIEW COMMITTEE MEMBERS</td>
<td>3</td>
</tr>
<tr>
<td>AOOS STAFF ATTENDING:</td>
<td>4</td>
</tr>
<tr>
<td>OBSERVERS</td>
<td>4</td>
</tr>
<tr>
<td>AOOS DATA MANAGEMENT REVIEW PROCESS METHODOLOGY</td>
<td>5</td>
</tr>
<tr>
<td>REVIEW CHARGE</td>
<td>5</td>
</tr>
<tr>
<td>REVIEW OBJECTIVES:</td>
<td>6</td>
</tr>
<tr>
<td>RECOMMENDATIONS</td>
<td>6</td>
</tr>
<tr>
<td>PRIORITY RECOMMENDATIONS</td>
<td>6</td>
</tr>
<tr>
<td>RECOMMENDATION 1: Establish a clear business and operations plan for the AOOS Data System</td>
<td>6</td>
</tr>
<tr>
<td>RECOMMENDATION 2: Implement a clear Intellectual Property policy</td>
<td>8</td>
</tr>
<tr>
<td>RECOMMENDATION 3: Create an operations model</td>
<td>9</td>
</tr>
<tr>
<td>RECOMMENDATION 4: Enhance trust in AOOS</td>
<td>11</td>
</tr>
<tr>
<td>RECOMMENDATION 4: Enhance trust in AOOS Error! Bookmark not defined.</td>
<td></td>
</tr>
<tr>
<td>ADDITIONAL RECOMMENDATIONS</td>
<td>12</td>
</tr>
<tr>
<td>RECOMMENDATION THEME: Discovery and documentation</td>
<td>12</td>
</tr>
<tr>
<td>RECOMMENDATION THEME: User support tools and strategies</td>
<td>13</td>
</tr>
<tr>
<td>RECOMMENDATION THEME: System usability</td>
<td>13</td>
</tr>
<tr>
<td>RECOMMENDATION THEME: Expand the use of the AOOS Data System to other programs</td>
<td>14</td>
</tr>
<tr>
<td>RECOMMENDATION THEME: Establish benefits and costs of contributing to regional/national/international efforts or programs</td>
<td>15</td>
</tr>
<tr>
<td>RECOMMENDATION THEME: Establish a method for establishing priorities for new products?</td>
<td>16</td>
</tr>
<tr>
<td>CONCLUSION</td>
<td>16</td>
</tr>
</tbody>
</table>
AOOS DATA MANAGEMENT REVIEW REPORT

INTRODUCTION

The Alaska Ocean Observing System (AOOS) was established in 2004 by a consortium of partner agencies and research institutions and serves as the Regional Association (RA) for integrating coastal and ocean observing activities in Alaska’s three regions: Arctic, Bering Sea/Aleutians and Gulf of Alaska. AOOS formally serves as the Alaska regional component of the U.S. Integrated Ocean Observing System (IOOS) and is now considered to be a Certified Regional Information Coordination Entity (RICE) under the authority of the Integrated Coastal and Ocean Observation System Act of 2009 (ICOOS Act).

As a member of IOOS, AOOS has a mandate to collect, organize, and provide access to Alaska oceanographic data. These data need to be easily understandable, electronically accessible and well organized to allow policy makers, industry and the general public to make well-informed decisions. To satisfy this mandate, AOOS supports a regional Data Assembly Center (DAC), operates a standards-based cyberinfrastructure, and provides a web-based data portal (called the Ocean Data Explorer) for the entire state of Alaska providing ocean, coastal and relevant watershed environmental data and information products. The collection of data, technology, policies and procedures, human resources, intellectual property and physical infrastructure employed to manage and use AOOS data is referred to as AOOS Data System (ADS). Data systems architecture, software engineering, and cyberinfrastructure operational services are provided by Axiom Data Science under contract to AOOS.

The Data Management Review Committee (henceforth the “Committee”) recognizes many strengths in the ADS. The AOOS team has strong engineering capabilities for data management, cyberinfrastructure development, and researcher support resulting in a well-engineered information technology system. The ADS adopts current best practices for system development to ensure a scalable foundation. The Axiom and AOOS teams are taking appropriate steps to lay a foundation for
reproducible and scalable science, while providing a platform to researchers that has the demonstrated potential to support cutting edge research. This platform is emerging as a nexus for both research and operational data product generation and distribution.

A major strength recognized is that the ADS is underpinned by scalable, high performance computing enabling advanced capabilities ranging from the monitoring of observing assets, to representing Indigenous knowledge, to advanced modeling. This results in an infrastructure that can grow and scale. It is clear that the ADS is continually evolving and is remaining at the forefront of cyberinfrastructure technologies, including the development of products that enable researchers to effectively manage their workflows and support transition into publishable datasets. There are Promising new developments such as the ship Automatic Identification System (AIS) data infrastructure/products, and the ability to create custom views of data using Data Dashboards. Scalability includes the demonstrated provision and adoption of community-recognized interoperability standards such as metadata (ISO 19115) and web services (e.g., OPeNDAP, SOS, WMS, WFS).

The strong working relationship between AOOS and Axiom is evident. Many Axiom staff have an academic or research background in similar fields to AOOS staff and researchers in the community which can help to ensure that they appropriately address AOOS staff and user needs.

The AOOS Board’s vision in supporting the development of AS as an AOOS flagship has enabled these developments and strengths. The Committee recommends continued investment in the sustainment and growth of this data management platform.

PARTICIPANTS

REVIEW COMMITTEE MEMBERS

Chair: Peter Pulsifer, NSIDC, IASC-SAON Arctic Data Committee

Sky Bristol, US Geological Survey

Jim Potemra, PacIOOS, Univ. of Hawaii
Derrick Snowden, IOOS Program Office

Dayne Broderson, UAF-GINA

Reyna Jenkyns, Ocean Networks Canada

**AOOS STAFF ATTENDING:**

Molly McCammon, Executive Director

Carol Janzen, Director of Operations

Holly Kent, Director of Administration and Outreach

**Axiom Data Science Staff Attending:**

Rob Bochenek, AOOS Data Manager, Axiom President/CEO

Stacey Buckelew, Axiom Director of Programs (partially, by phone)

Will Koeppen, Axiom Physical Scientist

Brian Stone, Axiom UI Developer

Jessica Austin, Axiom Software Engineer

Chris Turner, Axiom Data Librarian

**OBSERVERS**

Jennifer Dorton: SECOORA Observing System & Data Manager
AOOS DATA MANAGEMENT REVIEW
PROCESS METHODOLOGY

The AOOS Data Management Review process brought together an international, diverse group of experts to identify system strengths and weakness, assess usability, review the current customer service model with the goal of improving and enhancing the system over the coming decade. Additionally, the process examined the AOOS work plan and proposed resource allocation model.

Members of the review committee were invited based on their specific expertise in different areas of data management, experience working in the Arctic, or ability to provide feedback on task planning and funding. To ensure adequate background knowledge, prior to the review the Committee was provided with a broad range of different documents including but not limited to the 2016-2020 Strategic Operations Plan, Data Assembly Center and Data Management Subsystem Plan, various tasks and procedures documents, staff CVs and others.

A priority area for the review was system usability. In the weeks prior to the DMR meeting, the Committee was tasked with carrying out series of hands-on usability exercises and providing the AOOS team with feedback. Summary results of a Data Portal User Survey provided a broader perspective on user experience.

The Committee performed an on-site review with AOOS and Axiom staff in Anchorage from 28-30, November, 2017. Following a day of orientation presentations by Axiom staff on the 28th of November, the Committee met with AOOS staff on the 29th and 30th to deliberate and formulate the recommendations included in this report.

REVIEW CHARGE

Review Goal: To develop recommendations that will guide AOOS Data Management System development (including resource allocation for products and services) over the next decade.
REVIEW OBJECTIVES:

1. Provide committee assessment of system usability
2. Review customer service model and data management system resource allocation
3. Identify system strengths and weaknesses
4. Recommend areas for improvement and opportunities for program enhancement.

RECOMMENDATIONS

A series of recommendations are provided for consideration by AOOS staff, the AOOS Board, and in some cases, Axiom. The Committee felt that some recommendations are foundational and have a high priority while others are important but are less critical. Recommendations are presented as Priority and Additional.

PRIORITY RECOMMENDATIONS

RECOMMENDATION 1: ESTABLISH A CLEAR BUSINESS AND OPERATIONS PLAN FOR THE AOOS DATA SYSTEM.

Establishing a clear and detailed business and operations plan for the ADS is seen as foundational, urgent and of the highest priority. The results of Recommendations 2-4 are seen as part of business and operations, and more details can be found as part of the description of those recommendations.

AOOS staff should do preliminary research to establish the most appropriate method and form for establishing a plan. This review would include evaluation of different business and governance models that may be appropriate for optimizing the
effectiveness and efficiency of the relationship between AOOS and Axiom under existing or new working models.

A plan should include a clear statement of the vision, mission, goals, and high level objectives of the ADS. Beyond this, there is a need to clarify the functions of AOOS and Axiom, including a clear definition of AOOS and Axiom as “corporate” entities and their respective authorities, roles and responsibilities.

A risk analysis that identifies current and potential risks and vulnerabilities facing AOOS and Axiom under the current and candidate working models identified (see above) should be executed. The risk analysis will propose risk mitigation measures.

A clear statement of ownership, access, control and possession of ADS intellectual property and other jointly developed assets is needed. This would include a fair and adequate model for contributions of other partners such as other IOOS Regional Associations. For example direct or funded intellectual contributions of and to SECOORA and CeNCOOS need to be made clear along with their roles, rights, and responsibilities. See Recommendation 2.

A clear operations model should be developed that includes service level agreements, change management procedures, iterative software release model, and clarity on roles and responsibilities of partners. See Recommendation 3.

AOOS is a trusted organization and steward of data. There are a number of actions that can be taken to maintain and further enhance trust in AOOS and ADS. See Recommendation 4.

Lastly, the business and operations plan should include an outline and analysis of possible methods for attracting funding for products and activities where practical. A competitive and collaborative analysis should be included to help situate the ADS relative to other established or emerging systems and identify possible high value partners.
The Committee identified that the intellectual property (IP) being generated thought the development of ADS and related activities is not clearly identified, nor is there a clear and consistent policy or “regime” for managing IP (e.g. licensing, release of copyright agreements etc.).

Many software packages used in the ADS are developed by third parties and released under open source licenses, however there is lack of clarity with to the IP developed by Axiom and the IP established through the integration of packages to form the system as a whole. There is recognition that implementing a complex or onerous regime should be avoided to reduce the risk of stifling innovation. Axiom is interested in making any code developed freely available by default, however curating open source packages for public consumption to foster a user and contributor community requires resources that are not currently allocated by AOOS.

Regardless of the perceived usefulness of packages or the feasibility of curating open source software, the lack of a clear and consistent IP regime may put AOOS and/or Axiom at risk. IP law indicates that copyright (the tool used for software code) is retained by the creator unless released or managed through a license, employment agreement, or similar vehicle. To manage these risks, an IP regime should be established and used. Professional counsel should be engaged in developing this regime. At a conceptual level, development of such a regime should:

- Clearly identify the IP contained within the ADS;
- Identify roles and rights of IP stakeholders, particularly between AOOS and Axiom;
- Identify risks and mitigation measures related to current and possible IP management regimes
• Identify and possibly create an appropriate corporate management entity for IP (e.g. ADS as wholly or partly owned subsidiary);
• Establish a clear source code (IP) stewardship model. Includes defining code documentation, maintenance and release practices for different levels of code
• Confirm model for access to IP. For example, subscription vs. ownership for AOOS.

RECOMMENDATION 3. CREATE AN OPERATIONS MODEL

The current operations model has evolved organically over time. In some instances (e.g. prototyping, innovation) it is working well. As the system grows and "hardens," a more formal operations model is needed to ensure effectiveness, efficiency, and reliability. The model should include:

• A clear service level agreement, including which services (if any) would require and justify high service levels (e.g. 24/7) and expected response times
• Roles and responsibilities of internal and external human resources
• Detailed budget for different human and technical components of the system
• Software and product project management method(s) (e.g. could include adding a Product Owner to the AOOS team while implementing some form of "Agile" development process at Axiom). The plan would need to consider the cost and priority implications of such a model i.e. possible reallocation of funding from Axiom to AOOS
• A clear release process with room for open innovation and rapid development balanced with a clear product owner role to authorize feature and version release
• Management and balancing of research and production activities.
• Requirements for software and product documentation. In general, both should be well documented so that it can be shared as a template for other RAs etc.. Costs must be considered.
• Data and code lifecycle management
• Consider an explicit model for software modularization of features/products of the software from the back-end to the front-end such that they can be used independently rather than have deep ingrained interdependencies (e.g., AIS data handling system, research workspace, dashboards interface, etc.). Below is an example of what the Ocean Networks Canada executed some years ago to break out Oceans 2.0 functionality into core and extension modules (in concept, code, documentation, and related processes).

AOOS has established itself as a trustworthy, reliable organization. Continuing and enhancing trust is highly valuable and continued attention and resources should be applied to this priority. This has implications for growth of AOOS as well as resource allocation. As a starting point, the status and aspirations of AOOS as a trusted repository should be documented and made clearly visible to end user. This includes discussion of the RICE certification but also other possible certifications such as World Data System member certification or the newly emerging CoreTrustSeal (see https://www.coretrustseal.org/).

This process is an opportunity to align with requirements from funders and publishers to establish their requirements with respect to stewarding data in a trusted, certified data repository. This is increasingly a requirement for Principal Investigators to allow publication of peer-reviewed literature and/or fulfilling obligations under their funded project data management plans.

At a more detailed technical level, a formal approach for provenance tracking should be considered. Specifically, a method following the W3C PROV specification (https://www.w3.org/TR/prov-overview/) is recommended, as this has been demonstrated as a viable means within the Research Data Alliance community (https://www.rd-alliance.org/groups/research-data-provenance.html). Attention should be paid to recording meaningful provenance for the data caching/processing pipeline on real time data such that derived products and analytics show a clear PROV trace. The Research Workspace would also provide a good opportunity to both collect provenance for activities in the space and an avenue for tools to manage and curate provenance. This approach should include raw data and instrument maintenance records/files (e.g., calibration sheets) for AOOS operated or funded observing assets to support provenance, alternate usages, reprocessing capabilities. This may relate to the “Role” category that is included in the ISO 19115 metadata records for data. It is possible that there is some capacity already developed within new directions mentioned by Axiom around supporting field campaigns.
Continued development of the cyberinfrastructure and the demonstrated sensor data monitoring system and alert tools is highly valuable and should be continued. Having the ability to readily diagnose and intervene in data interruptions is critical to maintaining a high level of uptime and trust in the system.

**RECOMMENDATION 5. ANALYZE CURRENT AND POSSIBLE FUTURE STATE OF CLOUD TO CLOUD INTEROPERABILITY AND THE EMERGING “ECOSYSTEM” OF CLOUD PLATFORMS**

New Cloud-based infrastructures are emerging and as a result data are, theoretically, becoming more available to a wide range of different users. However, making data available on a cloud platform (e.g. Amazon Web Services, Microsoft Azure etc.), does not necessarily allow data to be easily shared across different data models and interfaces. Work is required to establish how the emerging ecosystem of cloud platforms will interoperate. This analysis should consider how the ADS can link to or ingest data from other clouds to enhance functionality, but also how ADS could serve data to other clouds. For example, the ADS cloud deployed in support for HPC modeling activities such as the NOAA Big Data initiative(s).

**ADDITIONAL RECOMMENDATIONS**

**RECOMMENDATION THEME: DISCOVERY AND DOCUMENTATION**

Don’t feel that you have to provide the only means of discovery through the ADF portals and applications. Let public search engines do some of your work for you (e.g., semantic mapping) by exposing the core discoverable units of the data system in optimized ways. This will require that each of these units have its own URL (not part of a “one-page app”) and include embedded structured metadata following the schema.org specifications. Not every single resource needs to be the point of entry from public search as those might just be search dead ends. Rather, think through the logical things you want people to find “in the wild” as entry points to the functionality you are providing.
Continue developing your embedding strategy where ADF components, tools, and visualizations can be embedded in other web sites. In many cases, you have mobilized and unchained data that are otherwise locked up and not very accessible from their original source. If you can find ways to feed that value-add back to the original sources that may sometimes be the most logical point of discovery and entry for users, this will get those users the value faster, build goodwill in the community, and bolster the AOOS brand.

**RECOMMENDATION THEME: USER SUPPORT TOOLS AND STRATEGIES**

Fully leverage the power of the technology platform that has been built where there are robust Application Programming Interfaces at the heart of the system. These enable rapid development of solutions that can be tailored for groups/types of users. When you find that you might have to write a tremendous amount of text to explain to users how to do something in the system, consider instead showing them interactively how to do it with a little bit of code or a simple API-driven interface.

Help enable your users to work for you and to be part of a vibrant community that is helping each other. Actively seek out cases where AOOS users have posted their own screencasts or other material demonstrating what they have done with your system. Highlight these on your web site and in presentations to help promote more of those contributions. Consider simple prizes and other low cost ways of engaging the community in providing this support.

Actively analyze the metrics you are now capturing in a cohesive way across ADF components to characterize and understand usage patterns and pathways. Where do users enter your systems, where do they end up, and what steps do they follow along the way? Developing an understanding of where users get stuck and what the primary modes of access are will help tailor both feature improvement and user support resourcing.

**RECOMMENDATION THEME: SYSTEM USABILITY**

In terms of the web tools, the software engineers are already building a more and more usable system by employing some of the best state of the art interface design patterns and established components that have proven successful in other communities. Continue following lean and uncluttered design principles; the less you have to explain, the more usable the system.
The bigger usability issues are at the level of data themselves. You can and should continue to address these through robust data documentation, quality assessment annotation, fitness for purpose analyses and reporting, and other metadata improvement methods. However, usage examples may provide the biggest improvement in overall usability. Exposing the software code behind every application and product you build as part of the core platform will provide one level of this as a natural byproduct of work you are doing anyway. Encouraging the community of users to share their code and scientific workflows via the Research Workspace, GitHub, or other venues will also contribute to a library of examples. Lightweight monitoring and curation of these community-generated resources will help you learn more about how the system is used and help guide users to the most useful content that promotes the usage patterns you would like to support. Continued development and publicising of high level APIs built on the lower level APIs (e.g., OGC services, etc.) that package common functionality into more efficient and attainable software tools for MatLab, Jupyter Notebooks, and other platforms will also be a powerful user enabler.

RECOMMENDATION THEME: EXPAND THE USE OF THE AOOS DATA SYSTEM TO OTHER PROGRAMS.

Expanding the use of the the ADS as practical and appropriate could provide many benefits. Specifically, there is efficiency in using a shared cyberinfrastructure. When designed with re-use in mind, technology funded by one sponsor project can be leveraged by other projects. Similarly, as the ADC user base grows, development and maintenance costs can be shared across projects. The diversity of features and functions that emerges from a larger and more diverse user base can act as a positive feedback -- new features and functions can make the platform more appealing to attract new users (contributors), further increasing efficiencies.

Continuing to move in this direction will require joint consideration (i.e. AOOS, other RAs, other contributors) of the technical and costs and benefits, but also the business, organizational and social aspects of possible or desired partners and what might impede adoption. As previously indicated, this should include a competitive and collaborative analysis, including potential users or partners in the private sector, civil society organizations and other.
RECOMMENDATION THEME: ESTABLISH BENEFITS AND COSTS OF CONTRIBUTING TO REGIONAL/NATIONAL/INTERNATIONAL EFFORTS OR PROGRAMS

Many regional, national and international initiatives are relevant to AOOS in different ways (scientific, operational, cyberinfrastructure etc.). There are opportunities to expand the reach of AOOS and the ADS through engagement in these initiatives. At the same time, engagement often requires resources of various kinds (e.g. developing technical interfaces, transforming data, attending meetings etc.), while the outcomes or benefits of engagement are not always clear. Moreover, can contributing to aggregation efforts such as portals be self-defeating by creating competition for the ADS metadata and data system? An analysis of potential partnerships with or engagement in the initiatives should be undertaken by AOOS. At minimum, this analysis would identify value propositions, a summary of the level of effort required to engage, and a strategic prioritization of engaging (or not) in a given initiative.

- Ocean Biogeographic Information System (OBIS) [http://www.iobis.org/](http://www.iobis.org/)
- International Oceanographic Data Exchange ([https://www.iode.org/](https://www.iode.org/))
- Interagency Arctic Research Policy Committee (IARPC) ([https://www.iarpccollaborations.org/index.html](https://www.iarpccollaborations.org/index.html))
- U.S. Arctic Observing Network Program
- Sustaining Arctic Observing Network ([https://www.arcticobserving.org/](https://www.arcticobserving.org/))
- Global Cryosphere Watch ([https://globalcryospherewatch.org/](https://globalcryospherewatch.org/))
- Glider Data Assembly Center ([https://gliders.ioos.us/](https://gliders.ioos.us/))
- Animal Telemetry Network DAC ([https://ioos.noaa.gov/project/atan/](https://ioos.noaa.gov/project/atan/))
RECOMMENDATION THEME: ESTABLISH A METHOD FOR ESTABLISHING PRIORITIES FOR NEW PRODUCTS?

The Committee discussed how to ensure that AOOS can establish priorities in the best possible way using one or more clear methods. Currently, decisions are based on local and expert knowledge of AOOS and AXIOM staff and partners. The committee recommends exploring different methods for priority setting. This could include the establishment of a “product committee” or a “product executive”, that would use a rubric to evaluate priority, relevance, value to science, urgency etc. and then balance these factors with resource availability. Screening new opportunities with a lean cost benefit analysis helps to quickly identify whether an opportunity is worth pursuing. Formal User Needs Analysis could also be a useful component. Any method implemented would need to link to the AOOS strategic plan and priorities, as well as consider various national priority plans - and cross-walking to those plans. The method should be applied regularly.

CONCLUSION

The panel applauds AOOS for conducting this review and opening up their investment to external scrutiny. The AOOS Data System is a remarkable set of capabilities that AOOS should be proud of. The skills assembled in the Axiom team are exemplary and capable of supporting the AOOS vision going forward. The recommendations provided in this report are intended to help AOOS make decisions on strategic investments over the next five years. They are heavy on process improvements and relatively light on technology changes as Axiom provides a state of the art technology foundation. Further, our recommendations focus on increasing the trust and business competitiveness of AOOS and are valid for a maturing system. Five years ago an external panel would likely have made very different recommendations. As AOOS seeks to further establish itself as a trusted repository of information and tool for research, management attention on reliability legal issues, business planning, intellectual property management, and strategic decision making will ensure that AOOS is poised to maintain and evolve their community standing as a reliable and innovative purveyor of ocean information. While considering our recommendations we recommend that AOOS seek to balance these new process improvements with maintaining the innovative culture and synergistic relationship with Axiom.