

# Toward Integrated Atlantic Ocean Observing Data Management: Industry Has a Role to Play Too\*

\*This article based in large part on remarks in the 18th Roger Revelle Commemorative (Distinguished) Lecture, National Academy of Sciences Ocean Studies Board, Washington, DC

Dawn J. Wright<sup>1,2</sup>

<sup>1</sup>Environmental Systems Research Institute (Esri), Redlands, California, United States of America

<sup>2</sup>College of Earth, Ocean, and Atmospheric Sciences, Oregon State University, Corvallis, Oregon, United States of America

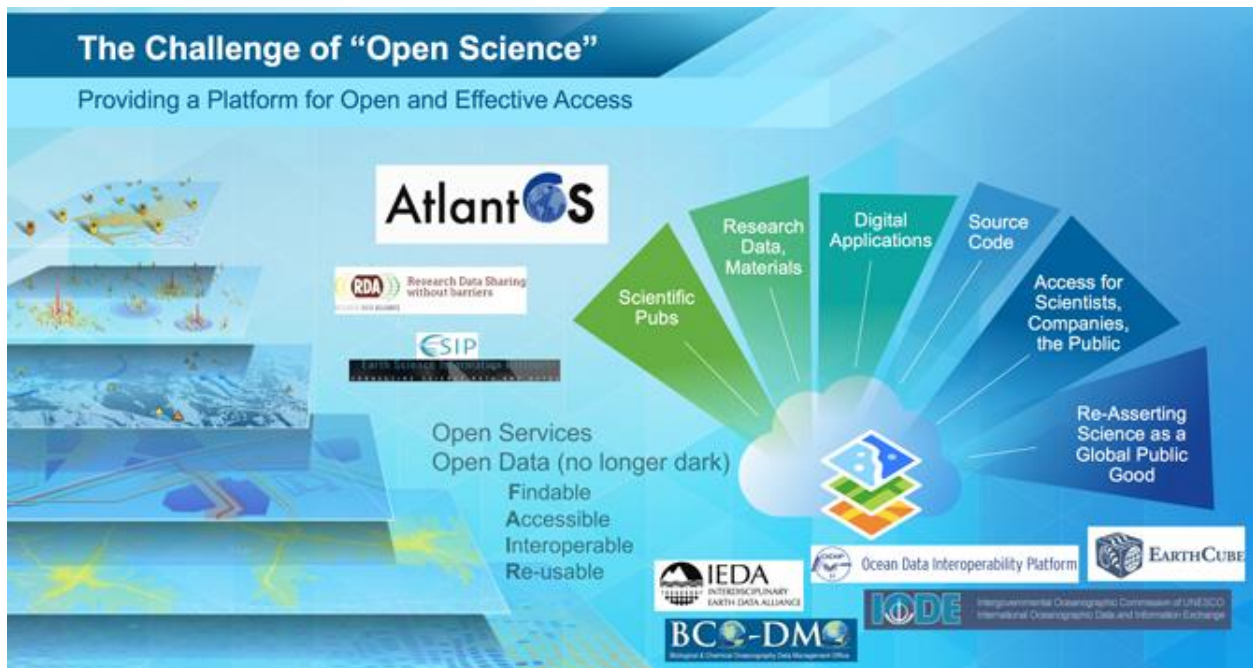
The variety or structural variability of data from ocean observing may be among the most compelling problems for the ocean science and management communities (e.g., Paolo et al., 2016). Data are coming from multiple sources and types (photos, video, audio, text, scientific observations, scientific models), multiple perspectives (governments, military, industry, non-governmental organizations or NGOs, etc.), which in turn have their various cultures for contributing and visualizing data. Although the number and type of ocean mapping applications continue to grow, there still exist overall inconsistencies in ocean data models, formats, standards, tools, services, and terminology. Hence, there is a need and provision of integrated data systems such as the one envisioned for Atlantic Ocean Observing System data within the framework of AtlantOS.

Tackling these problems has largely been in the realm of academia and federal agencies, but there is a new ocean data industry that is evolving to help meet these needs. It is often the private sector that is extremely skilled and efficient at providing the necessary sustainability and reliability of information via proven, well-engineered platforms, all with the necessary interoperability and openness as guaranteed through the adoption of established standards (Wright, 2017). It is estimated that: (1) 80% of the decision-making processes in ocean science and business depend on data collection, management, processing, and distribution; (2) accordingly, the data acquisition market is over \$80 billion, including ships, buoys, satellites, robots, ship-to-shore communications; and further (3) the data management market is estimated at \$5 billion, including software and associated costs (Rainer Sternfeld, PlanetOS, *pers. comm.*, April 23, 2013). This is fodder for effective public-private partnerships (PPPs) among academia, government, industry, and NGOs, especially when society is searching for sustainable solutions to multi-tiered environmental challenges.

It is in this spirit that the Environmental Systems Research Institute (aka Esri), a leading geographic information systems, data analytics, and data research and development company is pleased to partner with AtlantOS in the important areas of **data management, harmonization and integration**. The development of more effective integrated data management and interoperability infrastructure across all the sciences, including the ocean sciences, is a strategic pillar of the company, especially in its service to more than 350,000 organizations worldwide, including most of the world's governments on both sides of the Atlantic, and more than 7000 colleges and universities. The partnership will include free leveraging of Esri's cloud-based platform, ArcGIS Online, [www.arcgis.com](http://www.arcgis.com), where ~160 million maps requests are made by ~1.6 million users per day, 4-5 billion map tiles requests are made per month, and the platform has been subject to 94 billion map views since its inception circa 2012 (Wright et al., 2015).

Indeed, as compelling as big data (and small data) are, there is also the challenge of “dark data.” As aptly stated by Mascarelli (2009): “More and more often these days, a research project’s success is measured not just by the publications it produces, but also by the data it makes available to the wider community. Research cannot flourish if data are not preserved and made accessible. All concerned must act accordingly.” As discussed in the sections above, the massive amounts of data produced using modern digital technologies (including mapping technologies) has enormous potential for science and its applications in public policy, the non-profit sector, and business. But how should this deluge be shared and managed to support innovative and productive research that also reflects public values?

For the Atlantic Ocean AtlantOS is leading the way in fostering a data-centric “counter culture.” It is among the ecosystem of like-minded organizations such as the Research Data Alliance (RDA), the Federation of Earth Science Information Partners (ESIP), and specifically for the oceans community, the Intergovernmental Oceanographic Data and Information Exchange (IODE) of UNESCO’s Intergovernmental Oceanographic Commission, the Ocean Data Interoperability Platform, the Interdisciplinary Earth Data Alliance of Columbia University and the Biological & Chemical Oceanography Data Management Office of the Woods Hole Oceanographic Institution, the US National Science Foundation’s EarthCube initiative. For example, these organizations are developing best practices for fully cataloging and provisioning data using the same persistent identifiers in force for published papers, such as Digital Object Identifiers (DOIs). And an important vision for the future is that not only the tables, figures, statistics, and printed maps in published papers should be readily accessible, but the actual digital datasets themselves. This further pertains to not only data from the laboratory, but also to data collected in the field by various assets of ocean observing systems.



AtlantOS is among many multi-/trans-disciplinary organizations working toward open and effective access to data in the ocean sciences. Map layer graphic on right inspired by the six pillars of open science shared by Carthage Smith and Heidi Laine at Research Data Alliance Plenary 7.

Perhaps most importantly, AtlantOS and many other organizations are exercising the FAIR principle (Findable, Accessible, Interoperable, Re-usable) as part of several pillars of “open science” (e.g., Organisation for Economic Co-operation and Development, 2015). This is concerning the “what” (scientific publications, research data and materials, digital apps, source code), the “who” (scientists, companies, the public), and the “why” (re-asserting science as a global public good). And geospatial data and map industry partners such as Esri look forward to working closely with AtlantOS to move map data (i.e., geospatial data) from that of an underdeveloped or undervalued asset within an open data framework to that of more universally recognized and accepted data type such as the simple spreadsheet (as determined by the Civic Analytics Network (2017)). This will be particularly important in local, state, and national government circles where scientific data are used for public policy and on-the-ground decision making.

## References

- Civic Analytics Network. 2017. *An Open Letter to the Open Data Community*. <http://bit.ly/2m1Uvjo>.
- Mascarelli, A.L. 2009. Data's shameful neglect. *Nature* 461:145.
- Organisation for Economic Co-operation and Development. 2015. *Making Open Science a Reality*. OECD Science, Technology and Industry Policy Paper, OECD Publishing, Paris, 108 pp.
- Paolo, D., A. Leadbetter, and H. Glaves, eds. 2017. *Oceanographic and Marine Cross-Domain Data Management for Sustainable Development*. IGI Global, Hershey, PA, 425 pp.
- Wright, D.J. 2017. The role of aggregators, in Kruse, J., Crompvoets, J., and Perlman, F. (eds.), *GEOValue: The Socioeconomic Value of Geospatial Information*, Boca Raton, FL: CRC Press/Taylor & Francis Group, 117-128.
- Wright, D. J., Esri Ocean Team, and Esri Living Atlas Team. 2015. Should data frameworks be inherently multiscalar? A use case of the Living Atlas of the World. *Eos, Transactions of the American Geophysical Union* 96:IN13D-03.