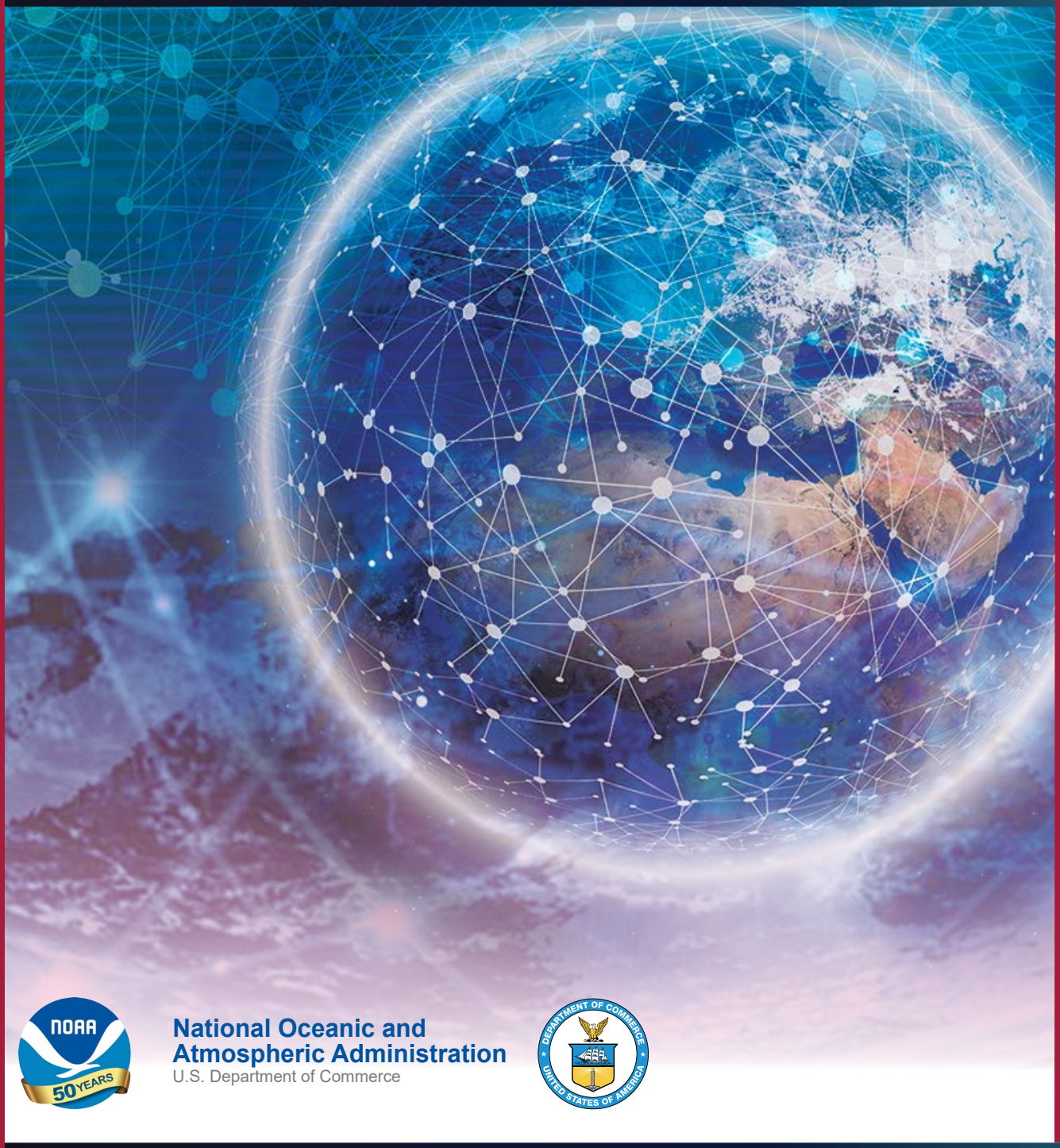
NOAA Data Strategy

Maximizing the Value of NOAA Data



National Oceanic and Atmospheric Administration
U.S. Department of Commerce

NOAA Science & Technology Focus Areas:

Unmanned Systems ▪ Artificial Intelligence ▪ 'Omics ▪ Cloud ▪ Citizen Science ▪ Data

July 2020

NOAA's Data Strategy

Maximizing the Value of NOAA Data

APPROVED:

Neil A. Jacobs, Ph.D.
Assistant Secretary of Commerce for Environmental Observation and Prediction
Performing the Duties of Under Secretary of Commerce for Oceans and Atmosphere

Tim Gallaudet, Ph.D., Rear Admiral, U.S. Navy (Ret.)
Assistant Secretary of Commerce for Oceans and Atmosphere / Deputy NOAA Administrator
STRATEGY SPONSOR

Stephen Volz, Ph.D.
Assistant Administrator
National Environmental Satellite, Data, and Information Service

Louis W. Uccellini, Ph.D.
Assistant Administrator
National Weather Service

Nicole LeBoeuf
Acting Assistant Administrator
National Ocean Service

Chris Oliver
Assistant Administrator
for Fisheries

Craig McLean
Assistant Administrator
Oceanic and Atmospheric Research

Rear Admiral Michael J. Silah
Director
Office of Marine and Aviation Operations and NOAA Corps

GOLDSTEIN.ZACHARY.G.1228698985 Digitally signed by GOLDSTEIN.ZACHARY.G.1228698985 Date: 2020.05.28 09:46:30 -04'00'

Zachary Goldstein
NOAA Chief Information Officer (CIO)
STRATEGY SPONSOR

July 2020

Executive Summary

Data are foundational to the mission of the National Oceanic and Atmospheric Administration (NOAA). Observations form the basis of NOAA's products and services; these are continuously collected from the bottom of the seafloor to the top of the atmosphere and beyond. These observations are used for a variety of activities such as analyses and predictions of the earth system, nautical charts, and support to decision making for safety of life and property, fisheries and ecosystem management, the Blue Economy[1], and any number of other NOAA missions. Each of these products and services, in turn, generate their own data, all of which must be effectively managed by the agency.

The purpose of the NOAA Data Strategy is to dramatically accelerate the use of data across the agency and with other key partners, maximize openness and transparency, deliver on mission, and steward resources while protecting quality, integrity, security, privacy, and confidentiality. The overall strategy is designed to serve as a framework for consistency that builds upon existing laws and regulations related to how NOAA uses and manages data, while being flexible and adaptable to external influences such as new policies, Executive Orders, stakeholder input, and new technologies that drive innovation within the agency.

The NOAA Data Strategy builds upon NOAA policy, practices, and processes related to data and how we manage it. The NOAA Data Strategy aligns with the June 4, 2019 released Office of Management and Budget (OMB) Memorandum, M-19-18 Federal Data Strategy - A Framework for Consistency[2], and the Federal Data Strategy's 2020 Action Plan[3]. It builds upon statutes and OMB information policy and guidance[4],with supplementary guidance on how agencies should manage and use Federal data. As with the Federal Data Strategy, the NOAA Data Strategy will constantly evolve to be regularly updated for consistency with new statutory or OMB information policy guidance. The NOAA Data Strategy is consistent with the Department of Commerce Strategic Plan for 2018-2022[5], the NOAA Information Resources Strategic Plan[6], and the NOAA 2020 Business Brief[7], and further codifies the NOAA Environmental Data Management Committee's Procedural Directives[8].

The NOAA Data Strategy will improve the management and overall value of the NOAA data enterprise when emerging science and technology have dramatically expanded the agency's data collections for the greater benefit to the agency, our partners, and the Nation. The NOAA Unmanned Systems, Artificial Intelligence (AI), 'Omics, and Cloud Strategies[9] will improve the coordination and use of emerging science and technology across the agency, and the NOAA Data Strategy is the foundation for each of these strategies. A fundamental consideration is that the goals and objectives within the NOAA Data Strategy directly support the other NOAA Strategies, thus requiring robust collaboration during implementation planning to develop a unified NOAA approach that supports the expanding application of emerging science and technology to effectively fulfill our mission priorities.

During the weeks of the HWT Experiment, forecasters assess a new tool using rapidly-updating high-resolution gridded Probabilistic Hazard Information (PHI) as the basis for next-generation severe weather warnings. The major emphasis of the HWT PHI experiment is on initial testing of concepts related to human-computer interaction while generating short-fused high-impact Probabilistic Hazard Information for severe weather. The long-term goal of this effort is to move the refined concepts and methodologies that result from this experiment into Hazard Services, the next generation warning tool for the NWS, for further testing and evaluation in the HWT prior to operational deployment.

Background

Data are at the heart of NOAA's $5 billion per year enterprise, and the NOAA Data Strategy provides a unified strategic approach for data governance that will deliver benefits across the organization. Each day, nearly every American relies on the data, products, and services NOAA provides. These products and services include daily weather forecasts, navigational tools to support the country's nearly $5.4 trillion in economic activity generated by U.S. seaports[10], assessments on the health of the nation's $200 billion fisheries[11], disaster response, and many more.

NOAA data are a critical strategic asset used to ensure accountability, manage operations, and to maintain and enhance the performance of the economy, public health, and welfare. The Nation already realizes the tremendous value of NOAA data[12]. Accurate and timely access to quality NOAA data significantly enhances the value and the return on the Nation's investment.

Deploying an ROV to install flow meters at the Middle Island Sinkhole in Lake Huron. (Photo Credit: David J Ruck/Great Lakes Outreach Media)

Strategic Imperative

Data are foundational to NOAA, underpinning virtually all aspects of our mission, including key NOAA priorities such as reducing the impacts of weather events and enabling the American Blue Economy. NOAA's data are also a vital national asset, serving the needs of countless stakeholders in industry, academia, other government agencies, and the public.

Aligning NOAA's capabilities with the constantly evolving needs of our stakeholders requires both collaboration and partnerships to deliver data and services in a way that stakeholders expect to consume them. Creative partnerships with commercial cloud providers allow NOAA to more effectively make its data publicly accessible. The NOAA Big Data Project[13] is a prime example. Demand for NOAA to store and process more than 20 terabytes of daily observational and model data means that innovative partnerships with cloud service providers are well suited to provide free public access to NOAA data.

NOAA's data holdings are vast, complex, and distributed, reflecting the diversity and footprint of NOAA's mission. Likewise, the systems and infrastructure that process, store, and disseminate NOAA's data are complex, distributed, and rarely shared across the agency. In the foreseeable future, the volume and velocity of our data are expected to increase exponentially with the advent of new observing systems and data-acquisition capabilities, placing a premium on our capacity and wherewithal to scale the information technology (IT) infrastructure and services to support this growth.

While there are existing policies, strategies, legislation, and Executive Orders that support the fundamental principles included in this NOAA Data Strategy, it is critical to emphasize that a consistent framework for data governance across the agency is effective and much more than just a compliance exercise. The adoption of this strategy will benefit the entire agency by:

- Stewarding and preserving valuable data assets, funded by U.S. taxpayers, for use by future generations;
- Aligning the value chain from observations through prediction and services, with sound data management principles and practices;
- Providing quality data and information required for decision makers;
- Ensuring readiness for artificial intelligence, machine learning, analytics, and other data science techniques that require ubiquitous access to structured, machine-readable data and metadata;
- Increasing NOAA customer satisfaction by enabling efficient access and use of our data;
- Reducing duplication of efforts and costs by improving efficient collaboration across NOAA entities, and through the use of common data practices and services;
- Enhancing NOAA workforce engagement through continuous improvements to data operations that enable NOAA employees to do their jobs more effectively;
- Modernizing NOAA's data processing through the use of big data, AI, machine learning, and other advanced analytical approaches; and
- Contributing to a strong economy by creating jobs in the information services and data science sectors.

The NOAA Data Strategy applies to all types of NOAA data including environmental, program, and statistical data, as well as mission-support data, including administrative, financial, performance, and workforce data.

The NOAA Data Strategy describes the vision, scope, goals, and objectives for establishing a foundation to effectively manage, share, and maximize the value of NOAA data assets. The NOAA Data Strategy goals and objectives are intended to guide NOAA agency-level implementation of the strategy. The NOAA Data Strategy applies to all types of NOAA data including environmental data, program data, and statistical data, as well as mission-support data including administrative, financial, performance, and workforce data. Data collected by external groups using NOAA funds (e.g. grant recipients, cooperative institutes) will continue to be subject to NOAA public access requirements[8,16]; however, these data would only be in-scope for the NOAA Data Strategy if they are stored in data systems managed by NOAA. The strategy does not define specific technologies or implementations.

Environmental data is the foundation of NOAA's mission. Through optimizing the availability and use of operational, research, and administrative data, NOAA is positioned to increase the value of data as a strategic asset in all facets of the organization. The overarching goal of the NOAA Data Strategy is to achieve readiness for tomorrow's data landscape.

Vision

A NOAA culture that values data as a strategic asset to understand the environment, create value for businesses, improve government efficiency, and provide quality services for the public.

Goals & Objectives

The NOAA Data Strategy includes a set of goals that are broadly applicable across NOAA mission areas, and supporting objectives for each that are intended to guide NOAA agency-level implementation of the strategy. The overarching goal of the NOAA Data Strategy is to achieve readiness for tomorrow's data landscape. NOAA maximizes the value of these assets by providing a consistent and transparent approach to NOAA data governance, stewardship, access, and use.

The goals and supporting objectives in this strategy are intended to directly improve the way NOAA manages its data and maximize the overall value of NOAA data assets to ensure the greatest benefit to the agency, our partners, and the Nation. Fundamentally, the objectives within and across the goals support and build upon each other, thus requiring robust collaboration during implementation planning and execution.

NOAA Data Strategy

Goal 1: Align data management leadership roles across the organization.

Goal 2: Govern and manage data strategically to most effectively steward the US taxpayers' investment.

Goal 3: Share data as openly and widely as possible to promote maximum utilization of NOAA data.

Goal 4: Promote data innovation and quality improvements to facilitate science and support data-driven decision making.

Goal 5: Engage stakeholders and leverage partnerships to maximize the value of NOAA data to the Nation.

GOALS	OBJECTIVES
1. Align data management leadership roles across the organization. When achieved, this goal will ensure there are sufficient authorities, roles, organizational structures, policies, and resources in place to transparently support the management and maximum use of NOAA data assets.	1.1 Empower the NOAA Chief Data Officer (CDO) to effectively represent data assets within the organization's planning and decision-making bodies. 1.2 Create a reporting and coordination structure throughout NOAA line and staff offices to effectively and collectively steward NOAA's data assets. Such a reporting structure could be analogous to the established intra-organizational means used to steward NOAA's financial or workforce assets. 1.3 Ensure that enterprise data governance, security, management, and stewardship are represented at NOAA's existing decision-making councils and Advisory Boards to facilitate effective data-driven decisions and investments.
2. Govern and manage data strategically to most effectively steward the US taxpayers' investment. When achieved, this goal will ensure that sufficient human and financial resources are available to support data-driven agency decision-making, accountability, and the ability to spur commercialization, innovation, and public use.	2.1 Establish a diverse Data Governance Body that has the authority to oversee NOAA's collective data operations, collectively govern how NOAA's data assets are managed, and provide strategic guidance for NOAA line and staff offices, relevant to their data and related activities, including Program Management, Acquisition, and Information Technology review boards. The Data Governance Body will report through NOAA's established strategic councils and executive governing bodies [i.e. NOAA Executive Panel (NEP) and NOAA Executive Council (NEC)]. 2.2 The Data Governance Body will review and consolidate existing NOAA data policies [e.g. NOAA Administrative Order (NAO) 212-15: Management of Environmental Data & Information[14], NAO 216-112: Policy on Partnerships in the Provision of Environmental Information[15], NOAA Procedural Directives[8], NOAA Public Access to Research Results (PARR)[16], and others]. 2.3 The Data Governance Body should facilitate the NOAA-wide use of common services to support NOAA data throughout the data lifecycle. 2.4 Adopt and use data standards to enable the mission, and share data using those standards.
3. Share data as openly and widely as possible to promote maximum utilization of NOAA data. When achieved, this goal will: - Promote equitable and appropriate access to NOAA data in open, machine-readable form and through multiple mechanisms, including both Federal and non-Federal providers, to meet stakeholder needs while protecting privacy, confidentiality, and proprietary interests; and - Disseminate NOAA data sets such that their authenticity is discoverable and verifiable by users throughout the information lifecycle, consistent with open data practices, and encourage appropriate attribution by users.	3.1 Develop a NOAA Open Data Plan, a Data Security Plan, and a collective Concept of Operations that describe how NOAA data can and will be shared appropriately inside NOAA, with other Federal Agencies, with NOAA's partners, and with the general public. They should incorporate not just mission and environmental data, but also administrative and business data necessary to evaluate NOAA's data practices. 3.2 Develop and maintain a comprehensive data inventory that accounts for all data assets created by, collected by, under the control or direction of, or maintained by NOAA. All items in the inventory must be documented with machine-readable metadata in a standard format, and published to the Federal data catalog. 3.3 Develop NOAA Data Licensing Guidance to ensure NOAA's data are by default "open" with no restrictions on their use or reuse, unless specifically otherwise restricted by law, regulation, or policy. 3.4 Share data as freely as possible to enable their use, while protecting privacy, confidentiality, and integrity. Adopt tiered data access practices to protect sensitive and confidential data and respect the rights of individuals and businesses. 3.5 Convey data authenticity and lineage to ensure that stakeholders can trust NOAA data, by verifying an authoritative source was used and preventing loss of data integrity.

4. Promote data innovation and quality improvements to facilitate science and support data-driven decision making. When achieved, this goal will: - Educate, empower, and continuously increase data literacy across NOAA's workforce by investing in training, tools, communities, and other opportunities to expand capacity for critical data-related activities such as analysis and evaluation, data management, and privacy protection; and - Encourage periodic review and optimization for the use of modern collaborative computing platforms to minimize costs, improve performance, and increase use.	4.1 As appropriate, implement enterprise data architecture solutions, linked with NOAA's Enterprise Architecture, to ensure that data and information systems are aligned to support NOAA's mission, science, and innovation. 4.2 Support innovative data solutions developed within NOAA so they can be leveraged by others within the agency and beyond. 4.3 Create workforce development and training opportunities to enable NOAA's scientists and data stewards to take full advantage of the available information and data science technologies. 4.4 Plan for the data infrastructure needs of new and expanding NOAA initiatives that will generate large amounts of data (e.g. AI, Unmanned Systems, 'Omics, and Citizen Science), taking a flexible approach as needed to account for the diverse and unique, project-specific characteristics of these data. 4.5. Establish clear and actionable guidelines for producing NOAA data in multiple open, machine-readable formats, including cloud-optimized and AI-ready file formats.
5. Engage stakeholders and leverage partnerships to maximize the value of NOAA data to the Nation. When achieved, this goal will: - Increase NOAA engagement with relevant agency partners, industry, academic, and other non-Federal users of data to share expert knowledge of data assets, promote wider use, improve usability and quality, and advance innovation and commercialization; and - Create and sustain partnerships that facilitate innovation with commercial, academic, and other partners to advance the agency's mission and maximize economic opportunities, intellectual value, and the public good.	5.1 Establish partnerships to enable effective and wide scaling of access to NOAA's data, the provision of expertise that supports the wider understanding of those data, and the effective use of NOAA information products by all. 5.2 Engage with NOAA's stakeholders to ensure that NOAA receives ongoing expert and timely feedback on NOAA's data practices, especially as they relate to the use of NOAA's data in the research and commercial sectors. 5.3 For NOAA data to meet stakeholder needs, the characteristics of each data set must be considered to maximize its value and optimal access. Common characteristics of data related to their value include data format, frequency of collection, quality, merchantability, and size of the data. The value of data for some applications also depends on their readiness for use relative to their collection time, such as forecasting high-impact weather. These types of characteristics inform the design of the data systems and delivery mechanisms. This necessitates user access to versatile configurations and systems.

Conclusion

The goals and objectives in this strategy will guide NOAA's future efforts to modernize our data practices, strengthening our environmental science and technology leadership for the coming decades. To ensure that the NOAA Data Strategy maximizes the value of our data assets, NOAA is developing a Data Strategy Implementation Plan or "Roadmap" that defines detailed action items, milestones, and responsibilities as resources permit. This will establish a modernized and unified NOAA data enterprise in support of the missions across NOAA, including close coordination with the NOAA Unmanned Systems, Artificial Intelligence, 'Omics, Cloud, and Citizen Science Strategies[9]. NOAA's data management activities will significantly improve the quality, access, and use of our public data assets and demonstrate our exceptional environmental science leadership.

Appendix A: Terminology

Data:
recorded information, regardless of form or the media on which the data are recorded. [*Reference: Evidence Act, II.a*[17]]

Data Asset:
a collection of data elements or data sets that may be grouped together [*Reference: Evidence Act, II.a*[17]]

Data Lifecycle / Information Lifecycle:
the stages through which information passes, typically characterized as creation or collection, processing, dissemination, use, storage, and disposition, to include destruction and deletion. [*Reference: OMB Circular A-130, Managing Information as a Strategic Resource*[18]]

Data Stewardship:
all activities that preserve and improve the information content, accessibility, and usability of data and metadata. [*Reference: NOAA Environmental Data Management Framework*[19]]

Data/Information Security:
the protection of information and information systems from unauthorized access, use, disclosure, disruption, modification, or destruction in order to provide:

- a) Integrity: guarding against improper information modification or destruction, including ensuring information non-repudiation and authenticity;
- b) Confidentiality: preserving authorized restrictions on access and disclosure, including means for protecting personal privacy and proprietary information; and
- c) Availability: ensuring timely and reliable access to and use of information (44 U.S.C. § 3552). [*Reference: OMB Circular A-130, Managing Information as a Strategic Resource*[18]]

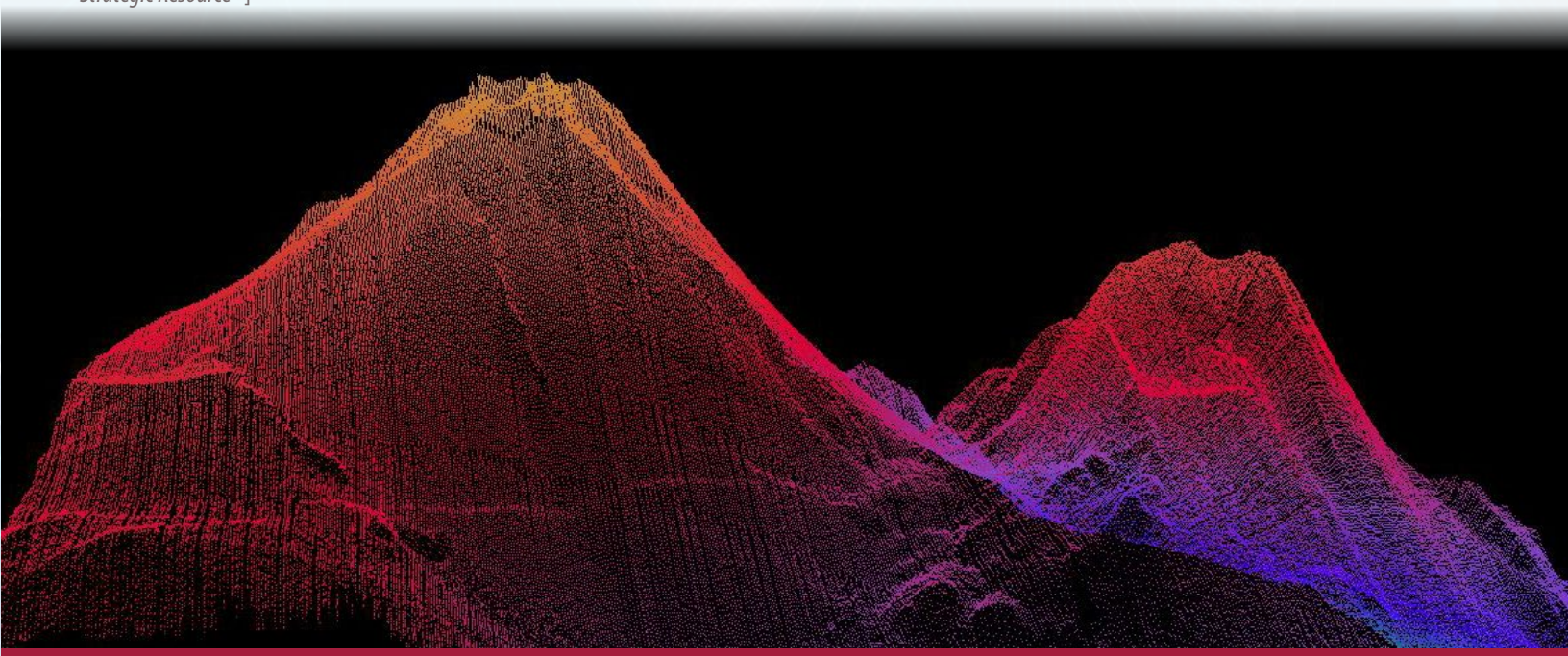
Metadata:
structural or descriptive information about data such as content, format, source, rights, accuracy, provenance, frequency, periodicity, granularity, publisher or responsible party, contact information, method of collection, and other descriptions. [*Reference: Evidence Act, II.a*[17]]

Machine-Readable:
in a format that can be easily processed by a computer without human intervention while ensuring no semantic meaning is lost. [*Reference: Evidence Act, II.a*[17]]

Public Data Asset:
a data asset, or part thereof, maintained by the Federal Government that has been, or may be, released to the public. [*Reference: Evidence Act, II.a*[17]]

Open Data:
a public data asset that is machine-readable; available (or could be made available) in an open format; not encumbered by restrictions, other than intellectual property rights, that would impede the use or reuse of such asset; and based on an underlying open standard that is maintained by a standards organization. [*Reference: Evidence Act, II.a*[17]]

Open License:
a legal guarantee that a data asset is made available at no cost to the public; and with no restrictions on copying, publishing, distributing, transmitting, citing, or adapting such asset. [*Reference: Evidence Act, II.a*[17]]



References

1. Leadership to Power the American Blue Economy, 2019. https://www.noaa.gov/sites/default/files/atoms/files/Leadership%20in%20Powering%20the%20American%20Blue%20Economy%20May%202020_0.pdf

2. OMB Memorandum M-19-18: Federal Data Strategy - A Framework for Consistency, June 2019. https://www.whitehouse.gov/wp-content/uploads/2019/06/M-19-18.pdf

3. Federal Data Strategy Year-1 (2020) Action Plan, December 2019. https://strategy.data.gov/assets/docs/2020-federal-data-strategy-action-plan.pdf

4. OMB Memorandum, M-13-13: Managing Government Information as an Asset throughout its Life Cycle to Promote Interoperability and Openness, May 2013. https://obamawhitehouse.archives.gov/sites/default/files/omb/memoranda/2013/m-13-13.pdf

5. Department of Commerce Strategic Plan 2018-2022, February 2018. https://www.commerce.gov/sites/default/files/us_department_of_commerce_2018-2022_strategic_plan.pdf

6. NOAA OCIO Information Resources Management Strategic Plan, May 2017. https://www.cio.noaa.gov/docs/NOAA-IRM-Strategic-Plan-Final-050117-without-Initiatives.pdf

7. 2020 NOAA Business Brief, April 2020. https://www.noaa.gov/sites/default/files/atoms/files/NOAA%202020%20Business%20Brief%20April%202020.pdf

8. NOAA Environmental Data Management Committee's Procedural Directives. https://nosc.noaa.gov/EDMC/PD.all.php

9. NOAA Science and Technology Focus Areas, 2020. https://nrc.noaa.gov/NOAA-Science-Technology-Focus-Areas

10. 2014 National Economic Impact of the U.S. Coastal Port System, American Association of Port Authorities, March 2015. Retrieved February 24, 2017, from https://www.aapa-ports.org/advocating/PRDetail.aspx?ItemNumber=22306

11. Fisheries Economics of the United States, 2014, NOAA National Marine Fisheries Service, 2016. https://repository.library.noaa.gov/view/noaa/11988

12.NOAA By The Numbers, July 2018. https://www.performance.noaa.gov/wp-content/uploads/NOAA-by-the-Numbers-Accessible-Version-Corrected-17-JUL-18.pdf

13. Big Data Program website. https://www.noaa.gov/organization/information-technology/big-data-program

14. NOAA Administrative Order 212-15: Management of Environmental Data & Information, November 2010. https://www.corporateservices.noaa.gov/ames/administrative_orders/chapter_212/212-15.html

15. NOAA Administrative Order 216-112: Policy on Partnerships in the Provision of Environmental Information, July 2007. https://www.corporateservices.noaa.gov/ames/administrative_orders/chapter_216/216-112.html

16. NOAA Plan for Increasing Public Access to Research Results (PARR), February 2015. https://repository.library.noaa.gov/view/noaa/10169

17. H.R.4174 - Foundations for Evidence-Based Policymaking Act, January 2019. https://www.congress.gov/bill/115th-congress/house-bill/4174/text

18. OMB Circular A-130 (revised), Managing Information as a Strategic Resource, 2016. https://obamawhitehouse.archives.gov/sites/default/files/omb/assets/OMB/circulars/a130/a130revised.pdf

19. NOAA Environmental Data Management (EDM) Framework, March 2013. https://nosc.noaa.gov/EDMC/framework.php

The NOAA NSSL Windsond researcher launch balloons that are smaller than normal weather balloons, with up to eight in the air at once. (Photo by Christiaan Patterson (OU CIMMS/NOAA NSSL))

Other Supporting Policies and Legislation

President's Management Agenda Cross-Agency Priority (CAP) Goal: Leveraging Data as a Strategic Asset, March 2018. https://www.performance.gov/CAP/leveragingdata/

OMB Memo M-19-23, Phase 1 Implementation of the Foundations for Evidence-Based Policymaking Act of 2018: Learning Agendas, Personnel, and Planning Guidance, July 2019. https://www.whitehouse.gov/wp-content/uploads/2019/07/M-19-23.pdf

Department of Commerce Guidance on Implementing the Evidence Act, March 2019.

Geospatial Data Act of 2018 (Subtitle F of the FAA Reauthorization Act). https://www.congress.gov/bill/115th-congress/house-bill/302

Executive Order on Maintaining American Leadership in Artificial Intelligence, The White House, 11 Feb 2019. https://www.whitehouse.gov/presidential-actions/executive-order-maintaining-american-leadership-artificial-intelligence/

NOAA Partnership Policy, https://www.noaa.gov/work-with-us/partnership-policy

Privacy Act of 1974, https://www.justice.gov/opcl/privacy-act-1974

Acknowledgements

The NOAA Data Strategy is led by Timothy Gallaudet, Ph.D., Rear Admiral, U.S. Navy (Ret.), Assistant Secretary of Commerce for Oceans and Atmosphere / Deputy NOAA Administrator. Kim Valentine from NOAA's National Ocean Service is coordinating the overall process in collaboration with the NOAA Data Strategy Team (below) and the NOAA Environmental Data Management Committee (EDMC), with executive sponsorship from the NOAA Observing Systems Council (NOSC) and the NOAA CIO Council.

National Environmental Satellite, Data, and Information Service (NESDIS)
Nancy Ritchey
Dana Ostrenga
Jason Cooper

National Marine Fisheries Service (NMFS)
Karen Sender
William (Bill) Michaels

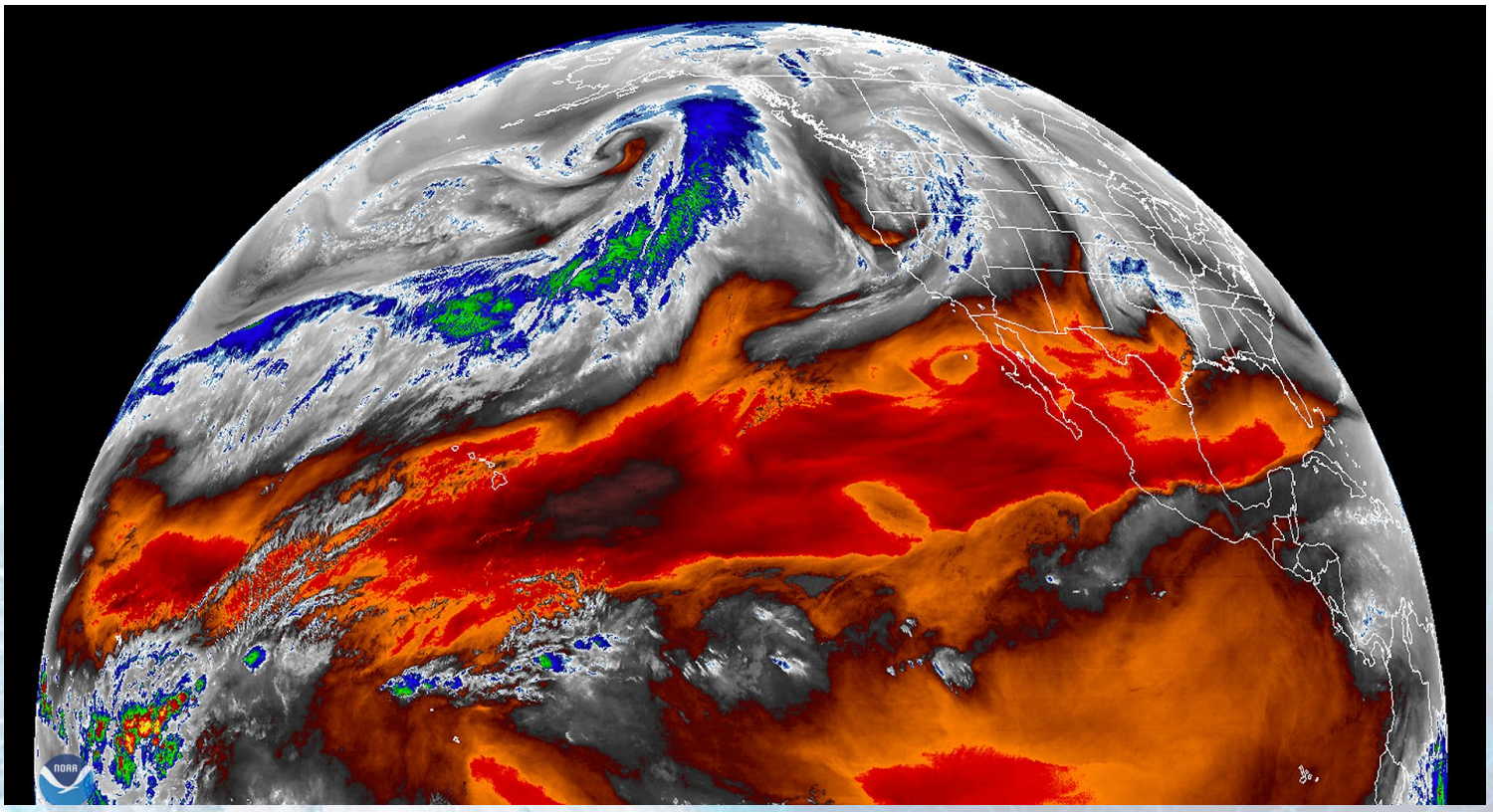
National Ocean Service (NOS)
Kim Valentine
Tyler Christensen

National Weather Service (NWS)
Jordan Gerth

Office of Oceanic and Atmospheric Research (OAR)
Eugene Burger
Lacey Mason

Office of Marine and Aviation Operations (OMAO)
Kevin Jackson
Solomon Tadele

NOAA Office of the CIO
Ed Kearns
Adrienne Simonson
Matt Austin



NOAA GOES-West satellite views atmospheric river in the Pacific Ocean.

All photos are courtesy of NOAA unless permission is expressly stated otherwise.