The Ocean Carbon and Acidification Data System

Li-Qing Jiang[1,2], Alex Kozyr[1,2], John Relph[2], Errol Ronje[3], Linus Kamb[4], Eugene Burger[4], Jonathan Myer[5], Liem Nguyen[6], Krisa M. Arzayus[7], Tim Boyer[2], Scott Cross[8], Hernan Garcia[2], Patrick Hogan[3], Kirsten Larsen[3], and A. Rost Parsons[3]

[1]Cooperative Institute for Satellite Earth System Studies, Earth System Science Interdisciplinary Center, University of Maryland, College Park, Maryland 20740, United States.

[2]NOAA/NESDIS National Centers for Environmental Information, Silver Spring, Maryland 20910, United States.

[3]NOAA/NESDIS National Centers for Environmental Information, Stennis Space Center, Mississippi 39529, United States.

[4]NOAA/OAR Pacific Marine Environmental Laboratory, Seattle, Washington 98115, United States.

[5]NOAA/NESDIS National Centers for Environmental Information, Asheville, North Carolina 28801, United States.

[6]Department of Computer Science, University of Maryland, College Park, Maryland 20740, United States.

[7]NOAA/NOS Integrated Ocean Observing System, Silver Spring, Maryland 20910, United States.

[8]NOAA/NESDIS National Centers for Environmental Information, Charleston, South Carolina 29412, United States

Corresponding author: Li-Qing Jiang (Liqing.Jiang@noaa.gov).

Abstract

The Ocean Carbon and Acidification Data System (OCADS) is a data management system at the National Oceanic and Atmospheric Administration (NOAA) National Centers for Environmental Information (NCEI). It manages a wide range of ocean carbon and acidification data, including chemical, physical, and biological observations collected from research vessels, ships of opportunity, and uncrewed platforms, along with laboratory experiment results, and model outputs. Additionally, OCADS serves as a repository for related Global Ocean Observing System (GOOS) biogeochemistry Essential Ocean Variables (EOVs), e.g., oxygen, nutrients, transient tracers, and stable isotopes. OCADS endeavors to be one of the world's leading providers of ocean carbon and acidification data, information, products, and services. To provide the best data management services to the ocean carbon and acidification research community, OCADS prioritizes adopting a customer-centric approach and gathering knowledge and expertise from the research community to improve its data management practices. OCADS aims to make all ocean carbon and acidification data accessible via a single portal, and welcomes submissions from around the world: https://www.ncei.noaa.gov/products/ocean-carbon-acidification-data-system/.

Introduction

With ~71% of its surface area covered by the ocean, the Earth is sometimes called a "Blue Planet'' or "Water Planet''. The ocean plays a critical role in climate regulation by sequestering and storing approximately 25% of anthropogenic carbon dioxide (CO_2), one of the primary greenhouse gases emitted by human activities [1-6]. Furthermore, the ocean absorbs about 90% of the excess heat trapped in the Earth system and transports heat from the equator to the poles[7,8]. Additionally, the ocean provides various ecosystem goods and services that are essential to our society. Billions of people, especially those living in the coastal communities, depend on the ocean for their food security[9,10], recreation[11], livelihood[12,13], and natural and cultural heritage[14].

Since the beginning of the Industrial Revolution around 1750, human activities have released ~2.5 trillion metric tons of carbon dioxide[1-6], resulting in an atmospheric CO_2 increase of ~50% (~419 parts per million, or ppm, 2022 annual average)[15] compared to the 1750 level of ~277 ppm[16], causing our climate to change[17,18]. Without the carbon capture and sequestration services provided by the ocean, the atmospheric CO_2 today would have been ~80 ppm higher than the current level[3].

This large influx of carbon dioxide is altering the ocean's chemistry. After CO_2 is absorbed by seawater, a portion of it will react with water to form carbonic acid, which will then dissociate and release hydrogen ions (H^+). The extra H^+ will associate with carbonate ions (CO_3^{2-}) to increase the concentration of bicarbonate ions (HCO_3^-). The net result is that the mildly alkaline ocean is becoming more acidic (and less alkaline), and its carbonate ion (CO_3^{2-}), a building block for many marine organisms, has been decreasing. This process is commonly referred to as "ocean acidification (OA)"[19-26]. Ocean acidification is making it harder for marine calcifiers (e.g., mollusks, crustaceans, coral) to build and maintain their shells and skeletal structures, endangering coral reefs and marine ecosystems more broadly[27-30]. It can also negatively affect non-calcifying organisms, e.g., disrupting the use of chemical sensations to find food or avoid predators for certain species, thus potentially shifting the food web, and changing the community composition and structure[31,32]. Even apex predators (e.g., large whales) may experience the effects of ocean acidification in the future as decreasing pH levels could change the acoustic characteristics of the ocean, thus alter or disrupt their communication and foraging behaviors as the ocean becomes noisier[33-35].

On timescales of decades to millennia, the ocean imposes a dominant control over atmospheric CO_2 levels, due to its vast size and efficient exchange of CO_2 with the atmosphere[5,36,37]. It contains the majority (~95%) of the total active pool of inorganic carbon in the Earth's surface, at ~38,000 billion tons of carbon, which is ~45 times that of the atmosphere[38-40]. Geoengineering efforts, such as marine carbon dioxide removal (mCDR), could potentially change the ocean's chemistry and impact the marine ecosystems[41]. Therefore, understanding carbon cycling in the ocean is critical for the research of global climate change, ocean acidification, marine CDR, and their downstream effects on the marine ecosystems, and the final impact on the human socio-economic structure[28,42].

Effective data management is a crucial aspect of the research efforts mentioned above[43, 44](Figure 1). After all, the ocean is a global system, and data collected from individual research cruises often need to be compiled into regional and global data products, before they can be used to support further oceanographic research, enable the Measurement, Reporting, and Verification (MRV) of mCDR for carbon credit accounting, and produce reports that can guide society's strategies for mitigating and adapting to environmental changes[45]. Data management provides an avenue where all data are (a) safeguarded for long-term access; (b) documented with common metadata and data standards and controlled vocabularies; and (c) findable and accessible. It plays a very important role in data reuse and valorization, model verification, and most importantly, quality control (QC), synthesis, and data product developments.

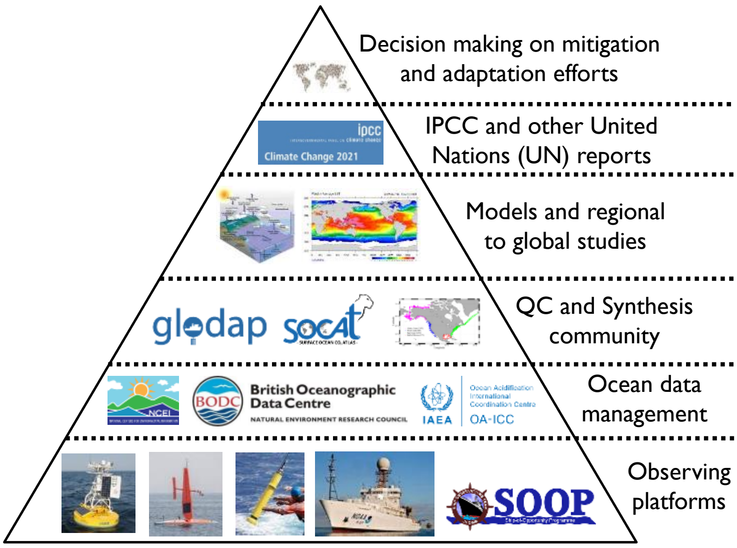

Figure 1. A schematic diagram showing the importance of ocean data management in promoting oceanographic research and product developments. (Credit: This figure is adapted from a diagram generated by Guidi et al.[46]).

The Ocean component of the Carbon Dioxide Information Analysis Center (CDIAC-Oceans, Oak Ridge National Laboratory, Tennessee, USA) played a very important role in the management of ocean carbon data from 1993 to 2016. Two factors contributed to the success of CDIAC-Oceans: its customer-centric approach, and the accessibility of all ocean carbon data collected by the global research community through a single data portal. However, CDIAC-Oceans' data holdings were not backed with a long-term archive. Its portfolio was focused on ocean carbon data and related hydrographic variables only. Its infrastructure was not designed to manage other types of OA data, such as physiological response biological studies. Worst of all, CDIAC-Oceans lost its funding from the Department of Energy (DOE) in 2016.

NCEI's OA data management efforts started with an OAP funded project called Ocean Acidification Data Stewardship (OADS) in 2012. By leveraging a modern community-driven rich metadata template[47] and NCEI's enterprise long-term archive management system with version control and controlled vocabulary management capabilities, OADS built state-of-the-art technical infrastructure that helped transform NCEI’s OA data management efforts from a top-down government task to a bottom-up community-driven service.

In November 2015, NCEI was notified that CDIAC-Oceans would lose its funding by the end of September 2016. To continue the service that was provided by CDIAC-Oceans, a new project called Ocean Carbon Data System (OCADS) was established at NCEI in January 2017, using the same technical infrastructure that was built by OADS. All CDIAC’s data holdings and webpages were transferred to NCEI, and the CDIAC-Oceans staff member was hired by NCEI as a cooperative institute employee to continue his work on the management of international ocean carbon data. The Ocean Carbon Data System could be considered as CDIAC-Oceans 2.0. All data holdings were upgraded with a much-improved modern metadata template[47]. Controlled vocabularies were applied to all aspects of the metadata to ensure accurate data findability. All data files were published into NCEI’s long-term archive with version control capabilities.

In September 2022, a new integrated system called the Ocean Carbon and Acidification Data System (OCADS) was formed at NCEI out of a merger of the former Ocean Acidification Data Stewardship Project (OADS) and Ocean Carbon Data System (OCADS). This merger was a response to feedback from the research community, who had encountered difficulties in choosing between the two systems for submitting and accessing data purposes. With the merger, the entire suite of ocean carbon and acidification data services offered by NCEI became available to the global research community, ushering in a new era of collaborative research possibilities. It's worth noting that the acronym OCADS remains the same, even though the new OCADS is a consolidation of the two previous systems.

Besides NCEI, current data management systems that have ocean carbon or acidification data in their portfolios include the CLIVAR and Carbon Hydrographic Data Office (CCHDO, U.S.A.), Biological and Chemical Oceanography Data Management Office (BCO-DMO, U.S.A.), the Ocean acidification International Coordination Center (OA-ICC, Monaco), PANGAEA (Germany), European Marine Observation and Data Network (EMODnet, Belgium), British Oceanographic Data Center (BODC, U.K.), Bjerknes Climate Data Centre (Norway), International Oceanographic Data and Information Exchange (IODE) Sustainable Development Goals (SDG) Portal (France), OceanOPS (France), and Joint European Research Infrastructure for Coastal Observatories (JERICO).

The Ocean Carbon and Acidification Data System has been the choice of data management services for most of the recent ocean carbon and acidification data product developments, e.g., the Surface Ocean CO_2 Atlas Version 2022 (SOCATv2022)[48], Global Ocean Data Analysis Product Version 2 (GLODAPv2.2022)[49], the Coastal Ocean Data Analysis Product in North America (CODAP-NA)[50], and the CARbon, tracer and ancillary data In the MEDiterranean sea (CARIMED).

OCADS excels in this arena because:

1. It is backed with a state-of-the-art long-term archive. All data are guaranteed to be available for at least 75 years, thanks to the NOAA records management requirements.
2. It has a stringent version control mechanism that ensures permanent

preservation of all historical versions.

3. OCADS provides a community-driven metadata display interface that is tailored to the needs and preferences of the oceanographic research community.
4. Controlled vocabularies are applied to all aspects of its data management to ensure accurate data findability.
5. OCADS manages a wide range of ocean carbon and acidification data, including chemical, physical, and biological observations, along with laboratory experiment results, and model outputs.
6. OCADS hosts one of the largest ocean carbon and acidification data repositories in the world, thanks to the data holdings transferred from CDIAC-Oceans.
7. OCADS has established mechanisms to support existing ocean carbon and acidification data product developments, e.g., SOCAT, and GLODAP.

Results

Mission

At its core, the mission of OCADS is to provide data management services that facilitate and support research on ocean carbon cycling and ocean acidification. This is accomplished through:

- Safeguarding their data in a well-supported federal archive to ensure long-term (≥75 years) access and version control,
- Serving as one of the world's leading providers of ocean carbon and acidification data, information, and products, and
- Providing data management support for quality control, synthesis, and data product development activities.

The OCADS data management services include dedicated support for acquiring data, publishing data into the archive, managing rich metadata with controlled vocabularies, and enabling online data retrieval and access. Another aspect of the OCADS portfolio is QCing data and developing coastal and open ocean data products[23,50,51]. In addition, OCADS has been playing a leading role in establishing numerous international standards for both metadata and data[47,52].

Data scope

OCADS manages a wide range of ocean carbon and acidification data (Table 1). Here, ocean carbon and acidification data are defined as data that contain at least one of these variables: carbon dioxide molecular ratio (xCO_2), pressure (pCO_2) or fugacity (fCO_2), total dissolved inorganic carbon content (DIC), total alkalinity content (TA), pH, hydrogen ion content ($[H^+]$), carbonate ion content ($[CO_3^{2-}]$), and calcium carbonate mineral saturation states for aragonite (Ω_{arag}) and calcite (Ω_{cal}). Besides these variables, OCADS also serves as a repository for other Global Ocean Observing System (GOOS) biogeochemistry Essential Ocean Variables (EOVs), i.e., oxygen, nutrients (e.g., Silicate, Phosphate, Nitrate), Transient Tracers (e.g., chlorofluorocarbons, or CFCs), nitrous oxide, particulate matter, stable carbon isotopes, and dissolved organic carbon[53].

Table 1. An inventory of the OCADS data holding by observation type as of January 15, 2023.

Observation type	Number of datasets
Profile/CTD	1173
Surface underway	822
Time-series	101
Laboratory experiment	29
Total	2230

Common types of data OCADS manages include:

- *In situ* observational ocean carbon and acidification data, including chemical, physical, and biological observations collected from research vessels, ships of opportunity, moorings, and other uncrewed platforms.
- Results from physiological response studies, including laboratory experiments, mesocosm studies, field experiments, and natural analogues.
- Model outputs.
- Data products (See Table 2 for examples)

Product development support

OCADS provides data management support for numerous ongoing projects focused on developing ocean carbon and acidification data products (Table 2). This service comprises archiving individual datasets from various cruises, producing tables with comprehensive lists of these datasets, and providing access to the developed data products. At present, OCADS can only provide support for data product development activities related to ocean carbon and acidification research due to limited resources.

Table 2. Major regional and global ocean carbon and acidification data product development efforts, with OCADS providing data management support.

Data product	Abbreviation	Area	Citation
The Surface Ocean CO_2 Atlas	SOCAT	Global	Bakker et al.[48]
The Global Ocean Data Analysis Project Version 2	GLODAPv2	Global	Lauvset et al.[49]
Coastal Ocean Data Analysis Product in North America	CODAP-NA	Coastal	Jiang et al.[50]
CARbon, tracer and ancillary data In the MEDiterranean sea	CARIMED	Regional	Álvarez et al.

SOCAT is a global data product that provides surface ocean fCO_2 measurements, primarily obtained from ships of opportunity (SOOP)[48]. It represents one of the most extensive collections of observational ocean carbon data. The latest release (SOCATv2022) contains 33.7 million fCO_2 values with an accuracy of better than 5 µatm. A further 6.4 million fCO_2 sensor data with an estimated accuracy of 5-10 µatm are available separately. GLODAP is a full water column open ocean data product, containing high-quality data from discrete bottle based measurements[49]. GLODAP covers 14 oceanographic variables, i.e., temperature, salinity, oxygen, nitrate, silicate, phosphate, DIC, TA, pH, chlorofluorocarbon (CFC-11), CFC-12, CFC-113, carbon tetrachloride (CCl_4), and Sulfur hexafluoride (SF_6). The most recent release, GLODAPv2.2022, includes measurements from close to 1.4 million water samples collected on 1085 cruises. CODAP-NA is an internally consistent data product for discrete inorganic carbon, oxygen, and nutrients on the North American ocean margins[50]. CODAP-NA's initial release (v2021) contains 3391 oceanographic profiles from 61 research cruises covering all continental shelves of North America, from Alaska to Mexico in the west and from Canada to the Caribbean in the east from 6 December 2003 to 22 November 2018.

Best practice data standards

Common data standards are crucial in facilitating future data utilization, particularly for quality control and synthesis efforts. Adhering to such standards can minimize uncertainties and errors that may result from ambiguous variable abbreviations, inconsistent quality control flags, and non-standardized calculations. A recently released best practice data standard for discrete chemical oceanographic observations[52] provides guidelines on various topics such as column header abbreviations, quality control flags, missing value indicators, standardized calculation methods for carbon system parameters, new tools for calculating thermodynamic variables using the International Thermodynamic Equations of Seawater 2010 (TEOS-10) equations[54], and new tools for calculating fCO_2 from dry-air mixing ratios.

Components

OCADS comprises three primary user interfaces that have been tailored to provide optimal data management support for research related to ocean carbon and acidification (Figure 2):

- The Scientific Data Information System (SDIS), a digital data submission interface,
- The Rich Metadata Management System (RMMS), a user-friendly metadata management and display interface, and
- The Ocean Carbon and Acidification Data System Portal (OCADS_portal), a data search and access interface.

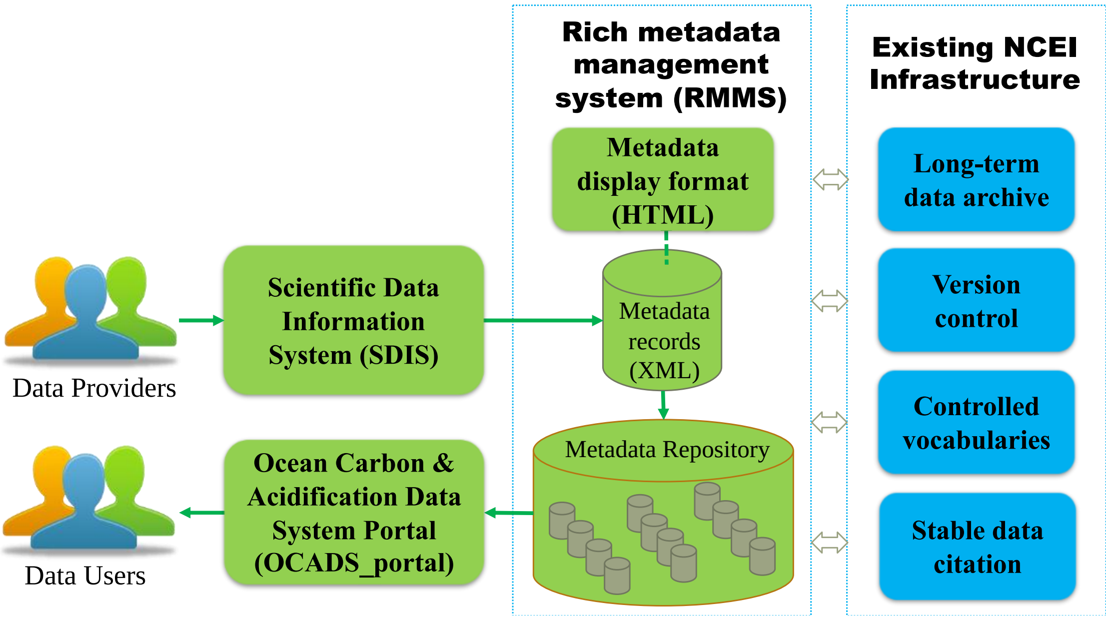

Figure 2. A schematic diagram showing the major components of the Ocean Carbon and Acidification Data System (OCADS).

The **Scientific Data Information System (SDIS)** is a digital data submission interface developed by NOAA's Pacific Marine Environmental Laboratory (PMEL) to streamline the process of submitting ocean carbon and acidification data to NCEI. (Figure 3). It enables a user to input metadata, upload data files, and submit the resultant data package to NCEI for review and eventual archival. The SDIS integrates rich metadata management capabilities that are designed to satisfy the research needs and requirements of the global ocean carbon and acidification research community[47]. The SDIS incorporates a user profile management system that permits data submitters to (a) maintain a record of all their prior submissions, (b) start a new submission by duplicating an existing record, and (c) save their work midway through a submission and resume later. In addition to new submissions, the SDIS also supports revision requests.

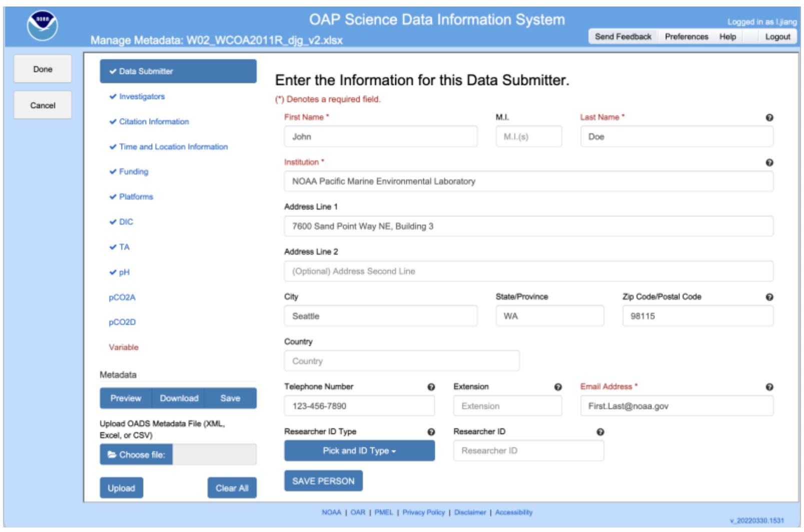

Figure 3. A screenshot showing the metadata interface of the Scientific Data Information System (SDIS, https://data.pmel.noaa.gov/sdig/oap/Dashboard/).

The **Rich Metadata Management System (RMMS)** is a metadata management and display interface that has been created to present the collected rich metadata information in a user-friendly format. At present, the RMMS is comprised of these sections: title, investigators, package description, data citation, identification information, types of study, a browse graphic, temporal coverage, spatial coverage, platforms, research projects, variable metadata sections, datasets related to this current dataset, and funding information[47]. The RMMS plays a crucial role in providing data management experiences that meet the needs and preferences of the global oceanographic research community.

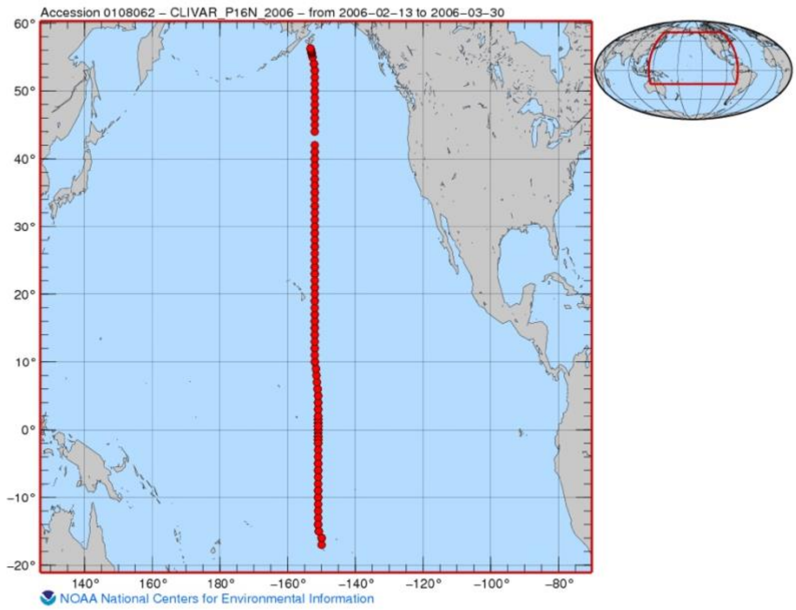

Figure 4. An example of a browse graphic from a dataset collected in the Pacific Ocean from 2006-02-13 to 2006-03-30 (Cruise_ID: CLIVAR_P16N_2006) (Credit: https://www.ncei.noaa.gov/data/oceans/ncei/ocads/metadata/0108062.html).

To ensure uniformity and ease of understanding, all OCADS dataset titles adhere to the template of "[observed properties] collected from [observation categories] using [instruments] from [research vessels or other platforms] in [sea names] during [research projects] from [start date] to [end date]. An example of an OCADS data title is: "*Dissolved inorganic carbon, total alkalinity, pH, temperature, salinity and other variables collected from profile and discrete sample observations using CTD, Niskin bottle, and other instruments from R/V Wecoma in the U.S. West Coast California Current System during the 2011 West Coast Ocean Acidification Cruise (WCOA2011) from 2011-08-12 to 2011-08-30*". Each OCADS dataset is accompanied by a browse graphic, providing users with a visual representation of the sampling locations or experiment setup, facilitating a quick understanding of the data. (Figure 4).

The "variable metadata section" is a one of the most important components of the RMMS, as it allows for the detailed documentation of all ancillary information related to a particular observed property, e.g., dissolved oxygen. This section can be repeated as many times as needed to document rich metadata information for all observed properties in the dataset. Information contained in this section will assist data users in comprehending the measurement details, such as the instruments used, methods applied, calibration procedures, and the associated data quality and uncertainty (accuracy/precision). Accuracy refers to how close the measurements are to the true or known values. Precision, on the other hand, refers to how close the measurements are to each other. Future improvements include the inclusion of a new metadata field that specifies whether a measurement is of weather or climate quality[55], as well as additional elements aimed at supporting marine carbon dioxide removal (CDR) research, e.g., types of alkalinization. An illustration of a variable metadata section is presented in Figure 5. Note that only the metadata elements that have been filled out by the data provider are shown, although the RMMS has the capacity to display all metadata elements as described in Table 2 of Jiang et al[47].

	bottle dissolved oxygen
Abbreviation:	OXYGEN_UMOL_KG
Unit:	micromol/kg
Observation type:	Discrete measurements from samples collected on CTD casts
In-situ / Manipulation / Response variable:	In-situ observation
Measured or calculated:	Measured
Sampling instrument:	Niskin bottle
Analyzing instrument:	Brinkman Dosimat automated titrator
Detailed sampling and analyzing information:	The analysis method is based upon the Carpenter (1965) whole flask titration of iodine, which is produced by an equivalent amount of dissolved oxygen. An automated titrator (Brinkman Dosimat) uses an amperometric end point detection as described by Culberson and Huang (1987) and modified for IBM-PC computers by Knapp et al. (1990). The nominal 125-ml iodine flasks are used for sampling are pre-calibrated so their volumes are precisely known. Samples were titrated within a few hours of being collected. 685 discrete oxygen samples were run to validate sensor O2 observations on the CTD package.
Quality flag convention:	OXYGEN_FLAG, WOCE quality control flags are used: 2 = good value, 3 = questionable value, 4 = bad value, 5 = value not reported, 6 = mean of replicate measurements, 9 = sample not drawn.
Researcher name:	Ann Swanson
Researcher institution:	Oregon State University, College of Earth, Ocean, and Atmospheric Sciences; PI: Burke Hales

Figure 5. An example of a "variable metadata section" from a dataset during the 2011 West Coast Ocean Acidification Cruise (https://www.ncei.noaa.gov/data/oceans/ncei/ocads/metadata/0123467.html).

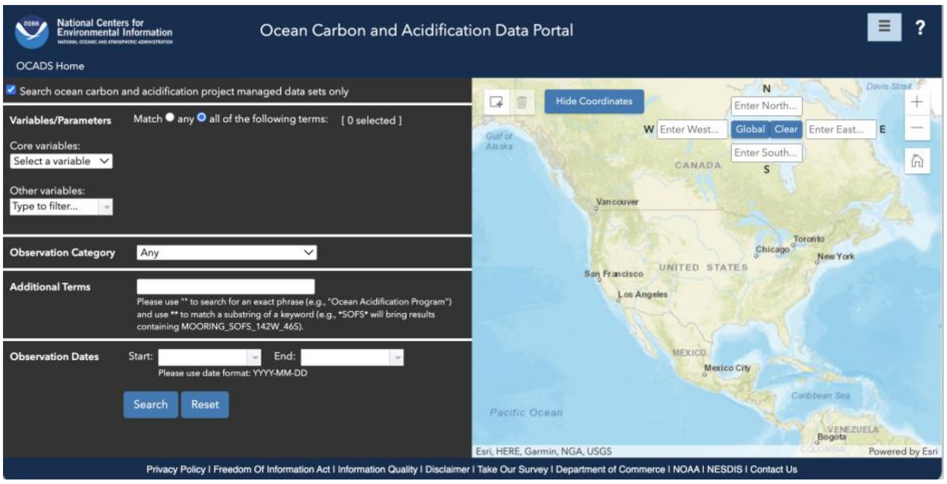

Figure 6. A screenshot showing the Ocean Carbon and Acidification Data System Portal (or OCADS_portal, https://www.ncei.noaa.gov/access/ocean-carbon-acidification-data-system-portal/).

The **Ocean Carbon and Acidification Data System Portal (OCADS_portal)** provides an interface for users to search and access all data holdings within OCADS

(Figure 6). It uses user-friendly technologies, such as auto-complete, dropdown menus, and OpenLayers maps, to enhance the user experience of searching for data. Currently, the Portal enables users to search for datasets based on five criteria:

- **Variables/Parameters:** The observed properties, e.g., water temperature, total alkalinity content, etc. A user can select one or multiple observed properties.
- **Observation Category:** How the observed properties are measured, e.g., surface underway, time-series, CTD profile, laboratory experiment, etc.
- **Additional Terms:** A powerful free-text search box allowing users to use any other elements of the metadata record to assist with the search, e.g., a research vessel name, the last name of an investigator, an expedition code (EXPOCODE), a cruise identifier (Cruise_ID), etc.[52]
- **Observation Dates:** The start and end dates of the observation. Only cruises with at least one data point collected during the period will show up in the search results.
- **Spatial Coverage:** A user can either input the bounding box information (longitudes and latitudes) or draw a rectangle on the map to define a geographic area that will constrain the spatial parameters of the query. Similarly, only cruises with at least one sampling station within the bounding box will show up in the search results.

Customers

OCADS recognizes the significance of (a) documenting all ocean carbon and acidification data with common metadata and data standards and controlled vocabularies, and (b) making them available through a single data portal. The program encourages the submission of data from across the globe and provides this service at no cost to all data producers, regardless of their location. In addition to data producers, OCADS also serves data users and consumers, including researchers, educators, decision-makers, private industry, and the general public from all nations.

Helpful tips

Here are some additional tips that will help users access data at NCEI:

- An “accession” refers to a dataset published at NCEI’s archives. The accession number is a 7-digit numerical number used to uniquely identify a dataset archived at NCEI.
- A “landing page” refers to NCEI’s generic metadata page, in contrast to the RMMS page that is served through OCADS. The former covers bare minimum elements such as title, abstract, citation, and keywords. The latter is much richer and is focused on providing the best metadata to meet the data needs of ocean carbon and acidification research.
- If only an accession number is available, below is how to access its metadata:
    - Link to the RMMS (available only if the dataset is published through OCADS, NNNNNNN is the accession number): https://www.ncei.noaa.gov/data/oceans/ncei/ocads/metadata/NNNNNNN.html

    - Link to NCEI’s landing page will be: https://accession.nodc.noaa.gov/NNNNNNN. This link works for all datasets within NCEI’s archives.
- If a user is taken to the landing page first by the search engine, here is how to access the rich metadata page: first click the “Documentation” tab, then click “Project metadata”.
- To see the version history of an Accession: first go to the landing page, then click the "Lineage” tab.

Links & Email

- **Ocean Carbon and Acidification Data System (OCADS):**
    - https://www.ncei.noaa.gov/products/ocean-carbon-acidification-data-system/.
- **Scientific Data Information System (SDIS):**
    - Link: https://data.pmel.noaa.gov/sdig/oap/Dashboard/.
    - Video tutorial: https://www.youtube.com/watch?v=ZZL_wQWr38A.
- **Rich Metadata Management System (RMMS):**
    - An example: https://www.ncei.noaa.gov/data/oceans/ncei/ocads/metadata/0123467.html
- **Ocean Carbon and Acidification Data System Portal (OCADS_portal):**
    - Link: https://www.ncei.noaa.gov/access/ocean-carbon-acidification-data-system-portal/.
    - Video tutorial: https://www.youtube.com/watch?v=DYFI0aH00FU.
- **OCADS contact:**
    - Email: noaa.ocads@noaa.gov.

Discussions

OCADS and international OA data management

OCADS operates within the Oceanographic Science and Development Branch (OSDB) of NCEI. As a component of the federal government, OCADS enjoys stable funding, which sustains its data management system. However, this does not imply that OCADS is biased towards NOAA or the United States. NCEI, the parent organization of OCADS, is a World Data Center for Oceanography designated by the International Council for Science (ICSU) through a resolution of the 29th United Nations General Assembly in 2008[56]. Like all other data that NCEI handles, 100% of the data that OCADS obtains will be made available to all users worldwide.

OCADS recognizes the importance of collaborating with other Data Analysis Centers (DACs) to achieve the common goal of offering top-notch data management services to the global oceanographic research community. Such partnerships not only help distribute the burden of international ocean carbon and acidification data management efforts, but also help overcome any regional or geopolitical barriers that could potentially prevent direct interaction between OCADS and local data producers. For DACs that need the support of a long-term archive, OCADS is eager to cooperate with them to transfer a copy of their data to NCEI's archive. This process is generally automated.

Ideally, all DACs, including OCADS, should function as regional nodes, aiding in the availability of all ocean carbon and acidification data via a centralized one-stop data portal. This can be accomplished by providing standardized metadata to the search engine of the agreed-upon one-stop portal. The most recent U.N. Ocean Acidification Research for Sustainability (OARS) data management initiative suggests designating the GOA-ON Portal as the envisioned one-stop OA data portal. Once this is implemented, users can employ the GOA-ON Portal to search for and retrieve all ocean carbon and acidification data of a specific type. After finding a dataset via the Portal, a user can then return to the respective regional DAC to access the data files and locate any pertinent metadata information.

For such a federated system to work properly, the participating DACs must fulfill the following minimum requirements:

1. Maintain a long-term archive to ensure uninterrupted data access into the future.
2. Provide strict version control capabilities, preserving all historical versions of a dataset on a permanent basis.
3. Utilize a community-driven common metadata template, e.g., the one used by OCADS, to collect comprehensive metadata information needed for ocean carbon and acidification research.
4. Present the collected metadata in:
    a. A user-friendly interface for metadata readability (e.g., HTML).
    b. A technical format to facilitate machine-to-machine interoperability (e.g., XML, SQL).
5. Employ common controlled vocabularies for successful data findability and easy machine-to-machine metadata exchange.
6. Support data citation with DOIs.
7. Establish a mechanism for sharing metadata with the agreed-upon one-stop portal, making the data searchable and accessible through the portal.

Before such a federated system is implemented, it is recommended that data producers share a copy of their data with OCADS to ensure timely inclusion in data products like SOCAT and GLODAP.

Data management vs. data product development

The concepts of "data management" and "data product development" are often conflated, but they are fundamentally different. It is true that they both involve working with data, but their similarities end there. Data management refers to the process of ingesting, storing, and disseminating data[43]. It includes tasks such as establishing long-term archives, maintaining version control, using metadata templates, adhering to data standards, and implementing controlled vocabularies. Data management is frequently mandated by government regulations and carried out by specialized national data centers.

Data product development, on the other hand, refers to the process of developing products and services (e.g., synthesis products, gridded climatologies, models, or predictions) out of data for the purpose of providing value to users. Unlike data management, data product development is rarely mandated by federal or local regulations. It can be carried out by anyone, including academic institutions.

Like other research activities, data product development also needs data management support.

Methods

OCADS was designed and built with the following rationales in mind:

(a) **A customer-driven service:** OCADS was designed to meet the data management needs and requirements of the global ocean carbon and acidification research community. We believe data management is a scientific effort that requires close collaboration with the research community. To ensure that our service is customer-driven, we rely on researchers to help us define many aspects of the OCADS data management, including metadata templates, controlled vocabularies, data standards, website design, and data submission and access interfaces.

(b) **Rich metadata management:** OCADS utilizes a community-driven rich metadata template[47] to collect detailed metadata information to enable data users, especially the synthesis community, to comprehend the measurement uncertainty and other sampling and calibration details, e.g., instrumentation, calibration, scales, units, biological subject, life stage, etc. of each variable. Rich metadata management is applied to the submission interface, metadata display interface, and the data search and access interface.

(c) **Controlled vocabularies:** Controlled vocabularies are lists of standardized terms. For example, some scientists could report their dissolved inorganic carbon measurements as “total dissolved inorganic carbon content”, while others may call it “dissolved inorganic carbon”, or “total carbon dioxide”. Multiple variations of terms for the same variable may cause confusion or decreased discoverability. Therefore, controlled vocabularies play a very important role to ensure successful data search. At OCADS, controlled vocabularies are applied to many groups of metadata elements, including observed properties (e.g., dissolved oxygen), observation types (e.g., surface underway, time-series), platforms (e.g., research vessels), sea names, instruments, institutions, people, countries, etc.

(d) **Stable data citation:** Each dataset archived at NCEI has an associated data citation. The citation contains information such as author list, title of the data, data repository, publication year, and an optional persistent identifier (e.g., DOI). DOI is strongly recommended for all OCADS datasets, although OCADS gives data providers the discretion of whether a DOI should be assigned or not. According to the NOAA Plan for increasing Public Access to Research Results (PARR)[57], “*NOAA National Data Centers are the only entities authorized to issue NOAA dataset identifiers*.”. This important guideline ensures that the same dataset will not be assigned multiple DOIs, potentially causing unnecessary confusions.

(e) **Long-term archive:** All OCADS datasets are archived using the NCEI enterprise archival infrastructure, ensuring the data will be available into the future (≥75 years). Besides a digital copy at NCEI’s server, all published datasets in NCEI’s archives are backed up to a staging disk on a backup server, as well as two offline tape-based repositories. Any data that are either new or changed are first backed up to the staging disk at a frequency of once per day, 7 days per week. Data on the staging disk is then migrated to tape copy #1 (on-premises). The same data are later copied to tape copy #2 (off-site).

(f) **Version control:** Unlike journal publications, datasets are often updated after they are published. This can happen after further QC is conducted or after additional data or metadata are gathered. While it is critical to provide future users with the latest version of the data, it is equally important to preserve all historical versions. Otherwise, research based on historical versions of the data can no longer be verified. In rare cases, the data may need to be reverted to a historical version. NCEI's enterprise archival infrastructure provides strict version control. After a dataset is published, no further changes can be made to the published version. Any changes to the data will trigger a new version of the dataset.

(g) **Preserving the original data:** In an era when the Findable, Accessible, Interoperable, and Reusable (FAIR) Data Principles[58] are often emphasized over traditional data formats like Microsoft Excel, OCADS recognizes the importance of preserving the original data. They often contain embedded QC comments, equations between different columns, color-coding, etc., which are critical to the future data use. Such information can often get lost during the conversion to FAIR-compatible data formats, e.g., NetCDF. Therefore, OCADS always ensures the preservation of the original data.

Data availability

All data presented in this article are available at https://www.ncei.noaa.gov/products/ocean-carbon-acidification-data-system/.

Code availability

Not applicable.

References

1. Sabine, C. L., Feely, R. A., Gruber, N., Key, R. M., Lee, K., et al. The oceanic sink for anthropogenic CO_2. *Science* **305,** 367-371, doi: 10.1126/science.1097403 (2004).
2. Gruber, N., Clement, D., Carter, B. R., Feely, R. A., van Heuven, S., et al. The oceanic sink for anthropogenic CO_2 from 1994 to 2007. *Science,* **363,** 1193-1199. doi: 10.1126/science.aau5153 (2019).
3. Friedlingstein, P., Jones, M. W., O'Sullivan, M., Andrew, R. M., Bakker, D. C. E., et al. Global carbon budget 2021. *Earth Syst. Sci. Data,* **14**, 1917-2005. https://doi.org/10.5194/essd-14-1917-2022 (2022).
4. Terhaar, J., Frölicher, T. L. & Joos, F. Observation-constrained estimates of the global ocean carbon sink from Earth System Models, *Biogeosciences,* **19**(18), 4431- 4457. https://doi.org/10.5194/bg-19-4431-2022 (2022).
5. DeVries, T. (2022). The Ocean Carbon Cycle. Annual Review of Environment and Resources, 47(1). https://doi.org/10.1146/annurev-environ-120920-111307.
6. Gruber, N., Bakker, D. C., DeVries, T., Gregor, L., Hauck, J., Landschützer, P., et al. Trends and variability in the Ocean Carbon Sink. *Nature Reviews Earth and Environment,* https://doi.org/10.1038/s43017-022-00381-x (2023).
7. Zanna, L., Khatiwala, S., Gregory, J. M., Ison, J. & Heimbach, P. Global reconstruction of Historical Ocean Heat Storage and Transport. *Proceedings of the National Academy of Sciences,* **116**(4), 1126–1131. https://doi.org/10.1073/pnas.1808838115 (2019).
8. von Schuckman, K., Cheng, L., Palmer, M. D., Hansen, J., Tassone, C., et al. Heat stored in the Earth system: where does the energy go? *Earth Syst. Sci. Data* **12,** 2013-2041, doi:10.5194/essd-12-2013-2020 (2020).
9. Costello, C., Cao, L., Gelcich, S., Cisneros-Mata, M. Á., Free, C. M., et al. The future of food from the sea. *Nature* **588,** 95–100, https://doi.org/10.1038/s41586-020-2616-y (2020).
10. Bennett, A., Basurto, X., Virdin, J., Lin, X., Betances, J., et al. Recognize fish as food in policy discourse and development funding. *Ambio* **50,** 981–989, doi:10.1007/s13280-020-01451-4 (2021).
11. Orams, M. B. The use of the sea for recreation and Tourism. *The GeoJournal Library* **78**, 161–173. https://doi.org/10.1007/978-1-4020-2780-2_9 (2004).
12. Charles, A. People, oceans and scale: Governance, livelihoods and climate change adaptation in Marine Social–Ecological Systems. *Current Opinion in Environmental Sustainability **4***(3), 351–357. https://doi.org/10.1016/j.cosust.2012.05.011 (2012).
13. Bari, A. Our Oceans and the Blue Economy: Opportunities and challenges. *Procedia Engineering* **194**, 5–11, doi: 10.1016/j.proeng.2017.08.109 (2017).
14. Ounanian, K., van Tatenhove, J. P. M., Hansen, C. J., Delaney, A. E., Bohnstedt, H., Azzopardi, E., et al. Conceptualizing coastal and maritime cultural heritage through communities of meaning and participation. *Ocean & Coastal Management **212***, 105806, doi:10.1016/j.ocecoaman.2021.105806 (2021).
15. Tans, P. & Keeling, R. Trends in atmospheric carbon dioxide. NOAA Global Monitoring Laboratory (https://gml.noaa.gov/ccgg/trends/) and Scripps Institution of Oceanography (https://scrippsco2.ucsd.edu/) (2023).
16. Meinshausen, M., Nicholls, Z. R., Lewis, J., Gidden, M. J., Vogel, E., et al. (2020).

The shared socio-economic pathway (SSP) greenhouse gas concentrations and their extensions to 2500. *Geoscientific Model Development* **13**(8), 3571–3605. https://doi.org/10.5194/gmd-13-3571-2020.

17. IPCC. *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L.Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)]. Cambridge University Press (2021).
18. IPCC. *Climate Change 2022: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* [H.-O. Pörtner, D.C. Roberts, M. Tignor, E.S. Poloczanska, K. Mintenbeck, A. Alegría, M. Craig, S. Langsdorf, S. Löschke, V. Möller, A. Okem, B. Rama (eds.)]. Cambridge University Press (2022).
19. Caldeira, K. & Wickett, M. E. Anthropogenic carbon and ocean pH. *Nature* **425**, 365–365. https://doi.org/10.1038/425365a (2003).
20. Feely, R. A., Sabine, C. L., Lee, K., Berelson, W., Kleypas, J., Fabry, V. J., et al. Impact of anthropogenic CO_2 on the $CaCO_3$ system in the oceans. *Science* **305**(5682), 362–366. https://doi.org/10.1126/science.1097329 (2004).
21. Orr, J. C., Fabry, V. J., Aumont, O., Bopp, L., Doney, S. C., Feely, R. A., et al. Anthropogenic ocean acidification over the twenty-first century and its impact on calcifying organisms. *Nature* **437**(7059), 681–686. https://doi.org/10.1038/nature04095 (2005).
22. Doney, S. C., Fabry, V. J., Feely, R. A. & Kleypas, J. A. Ocean acidification: the other CO_2 problem. *Annu. Rev. Mar. Sci.* **1,** 169–192, doi:10.1146/annurev.marine.010908.163834 (2009).
23. Jiang, L.-Q., Carter, B. R., Feely, R. A., Lauvset, S. & Olsen, A. Surface ocean pH and buffer capacity: past, present and future. *Sci. Rep.* **9,** 18624, doi:10.1038/s41598-019-55039-4 (2019).
24. Lauvset, S.K., Carter, B. R., Perez, F., Jiang, L.-Q., Feely, R. A., Velo, A., et al. Processes driving global interior ocean pH distribution. *Global Biogeochemical Cycles* **34,** e2019GB006229. doi:10.1029/2019GB006229 (2020).
25. Iida, Y., Takatani, Y., Kojima, A. & Ishii, M. Global trends of ocean CO_2 sink and ocean acidification: an observation-based reconstruction of surface ocean inorganic carbon variables. *Journal of Oceanography,* **77**(2), pp.323-358 (2021).
26. Hassoun, A.E.R., Bantelman, A., Canu, D., Comeau, S., Galdies, C., et al. Ocean acidification research in the Mediterranean Sea: Status, trends and next steps. *Frontiers in Marine Science,* p.1841 (2022).
27. Kawahata, H., Fujita, K., Iguchi, A., Inoue, M., Iwasaki, S., et al. Perspective on the response of marine calcifiers to global warming and ocean acidification—Behavior of corals and foraminifera in a high CO_2 world “hot house”. *Prog. Earth Planet Sci.* **6,** doi:10.1186/s40645-018-0239-9 (2019).
28. Doney, S. C., Busch, D. S., Cooley, S. R. & Kroeker, K. J. The impacts of ocean acidification on marine ecosystems and reliant human communities. *Ann. Rev. Environ. Resources* **45,** doi:10.1146/annurev-environ-012320-083019 (2020).
29. Gattuso J.-P., Magnan A., Billé R., Cheung W. W., Howes E. L., et al. Contrasting

futures for ocean and society from different anthropogenic CO_2 emissions scenarios. *Science* **349,** aac4722. 26138982 (2015).
30. Connell, S. D., Doubleday, Z. A., Foster, N. R., Hamlyn, S. B., Harley, C. D., Helmuth, B., et al. The duality of ocean acidification as a resource and a stressor. *Ecology* **99**, 1005-1010. https://doi.org/10.1002/ecy.2209 (2018).
31. Ashur, M. M., Johnston, N. K. & Dixson, D. L. Impacts of ocean acidification on sensory function in marine organisms. *Integrative and Comparative Biology* **57**(1), 63–80. https://doi.org/10.1093/icb/icx010 (2017).
32. Gunderson, A. R., Tsukimura, B. & Stillman, J. H. Indirect effects of global change: From physiological and behavioral mechanisms to ecological consequences. *Integrative and Comparative Biology,* **57**(1), 48–54. https://doi.org/10.1093/icb/icx056 (2017).
33. Hester, K. C., Peltzer, E. T., Kirkwood, W. J. & Brewer, P. G. Unanticipated consequences of ocean acidification: A noisier ocean at lower pH. *Geophysical Research Letters* **35**(19):1–5. doi:10.1029/2008gl034913 (2008).
34. Brewer, P. & Hester, K. Ocean acidification and the increasing transparency of the ocean to low-frequency sound. *Oceanography* **22**(4), 86–93. https://doi.org/10.5670/oceanog.2009.99 (2009).
35. Ilyina, T., Zeebe, R. E. & Brewer, P. G. Future ocean increasingly transparent to low-frequency sound owing to carbon dioxide emissions. *Nature Geoscience* **3**(1):18–22. doi:10.1038/NGEO719 (2010).
36. Revelle, R. & Suess, H. E. Carbon dioxide exchange between atmosphere and ocean and the question of an increase of atmospheric CO_2 during the past decades. *Tellus,* **9**(1), 18–27. https://doi.org/10.3402/tellusa.v9i1.9075 (1957).
37. Broecker, W. S. Glacial to interglacial changes in ocean chemistry. *Progress in Oceanography,* **11**(2), 151–197. https://doi.org/10.1016/0079-6611(82)90007-6 (1982).
38. Falkowski, P., Scholes, R. J., Boyle, E., Canadell, J., Canfield, D., et al. The global carbon cycle: A test of our knowledge of Earth as a system. *Science* **290**, 291-296, doi:10.1126/science.290.5490.291 (2000).
39. Ussiri, D.A. & Lal, R. The Global Carbon Inventory. In: *Carbon Sequestration for Climate Change Mitigation and Adaptation*. Springer, Cham. doi:10.1007/978-3-319-53845-7_4 (2017).
40. Schlesinger, W.H. & Bernhardt, E. *Biogeochemistry: an analysis of global change.* (4th ed.). Waltham, Mass.: Academic Press. ISBN: 9780128146088 (2020).
41. Burt, D. J., Fröb, F. & Ilyina, T. The sensitivity of the marine carbonate system to regional ocean alkalinity enhancement. *Frontiers in Climate,* **3**. https://doi.org/10.3389/fclim.2021.624075 (2021).
42. Falkenberg, L. J., Bellerby, R. G., Connell, S. D., Fleming, L. E., Maycock, B., Russell, B. D., et al. Ocean acidification and human health. *International Journal of Environmental Research and Public Health* **17**(12), 4563 (2020).
43. de La Beaujardière, J., Beegle-Krause, C. J., Bermudez, L., Hankin, S., Hazard, L., Howlett, E., et al. Ocean and Coastal Data Management. Proceedings of OceanObs'09: Sustained Ocean Observations and Information for Society. https://doi.org/10.5270/oceanobs09.cwp.22 (2010).
44. Brett, A., Leape, J., Abbott, M., Sakaguchi, H., Cao, L., Chand, K., et al. Ocean

data need a sea change to help navigate the warming world. *Nature* **582**, 181–183. doi: 10.1038/d41586-020-01668-z (2020).

45. Tanhua, T., McCurdy, A., Fischer, A., Appeltans, W., Bax, N., Currie, K., et al. What we have learned from the framework for ocean observing: evolution of the global ocean observing system. *Front. Mar. Sci.* **6**, 471. doi: 10.3389/fmars.2019.00471 (2019).
46. Guidi, L., Fernandez Guerra, A., Canchaya, C., Curry, E., Foglini, F., Irisson, J.-O., et al. Big Data in Marine Science. Alexander, B., Heymans, J. J., Muñiz Piniella, A., Kellett, P., Coopman, J. [Eds.] *Future Science Brief of the European Marine Board*, Ostend, Belgium. doi: 10.5281/zenodo.3755793 (2020).
47. Jiang, L.-Q., O'Connor, S. A., Arzayus, K. M. & Parsons, A. R. A metadata template for ocean acidification data, *Earth System Science Data* **7,** 117-125, doi:10.5194/essd-7-117-2015 (2015).
48. Bakker, D. C. E., Pfeil, B., Landa, C. S., Metzl, N., O'Brien, K. M., et al. A multi-decade record of high-quality fCO_2 data in version 3 of the Surface Ocean CO_2 Atlas (SOCAT). *Earth Syst. Sci. Data* **8,** 383–413, doi:10.5194/essd-8-383-2016 (2016).
49. Lauvset, S. K., Lange, N., Tanhua, T., Bittig, H. C., Olsen, A., et al. GLODAPv2.2022: the latest version of the global interior ocean biogeochemical data product, *Earth Syst. Sci. Data* **14**, 5543–5572, https://doi.org/10.5194/essd-14-5543-2022 (2022).
50. Jiang, L.-Q., Feely, R. A., Wanninkhof, R., Greeley, D., Barbero, L., et al. Coastal Ocean Data Analysis Product in North America (CODAP-NA) – an internally consistent data product for discrete inorganic carbon, oxygen, and nutrients on the North American ocean margins. *Earth Syst. Sci. Data* **13,** 2777–2799, doi:10.5194/essd-13-2777-2021 (2021).
51. Jiang, L.-Q., Feely, R. A., Carter, B. R., Greeley, D. J., Gledhill, D. K. & Arzayus, K. M. Climatological distribution of aragonite saturation state in the global oceans. *Global Biogeochemical Cycles* **29**, 1656-1673. doi:10.1002/2015GB005198 (2015).
52. Jiang, L.-Q., Pierrot, D., Wanninkhof, R., Feely, R. A., Tilbrook, B., et al. Best Practice Data Standards for Discrete Chemical Oceanographic Observations. *Front. Mar. Sci.* **8**, 705638. doi:10.3389/fmars.2021.705638 (2022).
53. Moltmann, T., Turton, J., Zhang, H.-M., Nolan, G., Gouldman, C., et al. A global ocean observing system (GOOS), delivered through enhanced collaboration across regions, communities, and New Technologies. *Frontiers in Marine Science,* **6.** https://doi.org/10.3389/fmars.2019.00291 (2019).
54. Intergovernmental Oceanographic Commission (IOC), Scientific Committee on Oceanic Research (SCOR) & International Association for the Physical Sciences of the Oceans (IAPSO). *The International Thermodynamic Equation of Seawater - 2010: Calculation and use of thermodynamic properties*. Intergovernmental Oceanographic Commission, Manuals and Guides No. 56, UNESCO (English), 196 pp (2010).
55. Newton J. A., Feely R. A., Jewett E. B., Williamson P. & Mathis J. *Global Ocean Acidification Observing Network: Requirements and Governance Plan*. *Second Edition*. Global Ocean Acidification Observing Network (GOA-ON). https://archimer.ifremer.fr/doc/00651/76343/ (2015).

56. Reed G., Rickards, L. & Troisi, A. Providing quality-assessed oceanographic data and services to the international science community. The First ICSU World Data System Conference, Kyoto, Japan, 3-6 September 2011. https://oceanexpert.org/downloadFile/13826 (2011).
57. NOAA Research Council. NOAA Plan for Increasing Public Access to Research Results: a response to the White House Office of Science and Technology Policy memorandum 'Increasing access to the results of Federal funded scientific research' issued February 22, 2013. Editor(s): De La Beaujardière, J., and Kaske, N. K., http://doi.org/10.7289/V5F47M2H (2015).
58. Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J. J., Appleton, G., Axton, M., et al. The Fair Guiding Principles for Scientific Data Management and Stewardship. *Scientific Data* **3**, https://doi.org/10.1038/sdata.2016.18 (2016).

Acknowledgements

OCADS is funded by the NOAA Ocean Acidification Program (OAP, Project ID: 21047), the NOAA Global Ocean Monitoring and Observing (GOMO) Program, and the NOAA National Centers for Environmental Information (NCEI). We are grateful to Dr. Libby Jewett and Dr. Dwight K. Gledhill of OAP for their initial funding in starting NCEI's OA data management effort, and for their continuous support since 2012. Our thanks also go to Kathy Tedesco and David Legler for their funding that enabled us to transfer data and web pages from CDIAC-Oceans to NCEI, and for their continued funding support since 2022. L-QJ and AK were supported by NOAA grant NA19NES4320002 (Cooperative Institute for Satellite Earth System Studies -CISESS) at the Earth System Science Interdisciplinary Center, University of Maryland.

We thank Joseph (Jack) Rhodes and Glen Reid of NCEI for help establish the current Ocean Carbon and Acidification Data System Portal. We thank Heather McCullough of NCEI for helping create the first version of the Drupal OCADS website. We thank John CartWright of NCEI (retired) for helping create the Mooring Data Viewer. We thank Mike Kruk and Rich Baldwin of NCEI for helping with multiple IT aspects of OCADS. We thank George Sinclair and Ryan Nelson of NCEI for explaining the details of NCEI’s back up process. We are grateful to Dr. Dwight K. Gledhill, Dr. Erica Ombres, and Kaitlin Goldsmith of OAP for their many constructive comments over the years. Most importantly, we want to thank all the OCADS data users and data providers whose comments help make OCADS a better system.

Author contributions

L-QJ created the first draft of the paper, and all authors contributed to the writing of the manuscript. L-QJ designed and coordinated the development of many OCADS components, including the metadata template, the Rich Metadata Management System, its underlying XML, XSL, and HTML formats, the OCADS_portal, the SDIS, main pages of the OCADS website, and the international data standard for discrete bottle-based data. L-QJ also designed and led the efforts to transfer a total of 2,296 data sets serving from 1,193 FTP directories at cdiac.ornl.gov (corresponding to 1,190 NCEI accessions) from the Carbon Dioxide Information Analysis Center (CDIAC-Oceans) to NCEI. OCADS was built on the success of the CDIAC-Oceans, which was first launched in 1993 by AK. AK processed the majority of the OCADS accessions, including all CDIAC-Oceans datasets that were later transferred to OCADS. AK helped review the design of the new OCADS website and created the many data access pages including the clickable maps and numerical data packages brought to NCEI from the former CDIAC-Oceans. AK also contributed tremendously to the transfer of data and webpages to NCEI. JR helped build the RMMS, the backend automation program for the SDIS, the automation programs to transfer data from CDIAC-Oceans to NCEI, along with many other IT aspects of OCADS. ER tested SDIS after its initial deployment and provided comments that improved the SDIS archive appraisal and ingest process and other aspects of OCADS. ER has been the subject matter expert working on the flow of biological OA data. LK and EB developed the Scientific Data Information System. JM created the Drupal portion of the OCADS website. LN wrote the code for the html pages of the new OCADS website, and the new OAP research project management system. KMA, TB, SC, HG, PH, KL and ARP (ranked alphabetically

based on their last names) provided resources and guidance for many aspects of OCADS.

Competing interests

The authors declare no conflict of interest.