SUPPORTING STATEMENT
NODC Send2NODC (S2N) Web Application
OMB CONTROL NO. 0648-0024

A. JUSTIFICATION

This request is for reinstatement with change of a previously approved information collection, with web-based information submission now the main method of submission.

Those who had been responsible for maintaining the PRA information collection several years ago (when this was still in paper form) retired and this task effectively 'fell through the cracks', with no internal transfer of responsibility for maintaining OMB approval. When we began the S2N development process, we came to understand that we needed to go through the OMB reinstatement process. NODC wants to get back on the proper path for regular PRA review, approval and renewals as we go forward with a better documented internal process and delegation of responsibility for the task.

1. Explain the circumstances that make the collection of information necessary.

The National Oceanic and Atmospheric Administration's (NOAA's) National Oceanographic Data Center (NODC) "Send2NODC" or "S2N" web application is a modernized, web-enabled representation of basic descriptive metadata for oceanographic observation data that was embodied in the analog (paper) NOAA Form 24-13 "Data Documentation Form" or "DDF". In 1994, Executive Order 12906 established the National Spatial Data Infrastructure and directed each agency to "document all new geospatial data it collects or produces… using the standard under development by the FGDC...". In 2010, the Federal Geographic Data Committee (FGDC) adopted the International Organization for Standardization (ISO) 19115/19139 suite of standards for use by Federal agencies to document geospatial data.

The voluntary information collection using the S2N web application or other representations of the DDF facilitates NOAA and others in Federal, State, or local governments, academic institutions, non-profit organizations, and others to provide clear descriptive information (metadata) about marine data. The information requested in S2N identifies the data provider and data collector(s), dates that data were collected, and other information that make marine data useful to other researchers and future users of the described data.

The S2N web application reorganizes descriptive information collected into an XML file that conforms to the internationally recognized and FGDC-approved ISO19115-2 Geospatial Metadata standard and the ISO19139 Geographic information Metadata XML schema implementation. A significant advantage for a data provider using S2N is that the information is collected through a series of short form pages that provide suggestions for entries from NODC-provided standard responses. The data provider may opt to use suggested entries or to enter free text in the form spaces.

2. Explain how, by whom, how frequently, and for what purpose the information will be used. If the information collected will be disseminated to the public or used to support information that will be disseminated to the public, then explain how the collection complies with all applicable Information Quality Guidelines.

NODC is the designated long term United States (U.S) repository for marine geospatial data and information and an internationally designated World Data System repository for sharing marine data. The information collected by S2N is used by NODC in the processes of acquiring, processing, quality controlling, and archiving marine geospatial data. The information collected by S2N may also be distributed with those related marine data to the community of users to enable them to better understand and make use of those data.
After review and acceptance for the repository by NODC staff, scientists and others use the geospatial metadata collected using S2N to discover the existence of marine data available for additional scientific re-use through web portals, geospatial clearinghouses (e.g., data.gov), and other web services. The S2N information collection web form facilitates creating suitable metadata to enable data discovery in these portals by collecting specific characteristics such as the dates during which data were collected, the area where data were collected (both by sea or sea region name and by latitude/longitude bounding coordinates), and ships/platforms used to collect marine geospatial data or information.

Specific characteristics (e.g., variable/data type, unit of measure, quality flags, etc.) are extremely important documentation that enable future users to fully understand and use marine and other geospatial data. By using S2N, this information may be collected at a relatively high level of specificity, with consistent representation of similar information across multiple information collection instances (submissions).

Information collection about the data provider (e.g. name, address, institutional affiliation) is used by NODC to establish communications with the data provider in the event there are questions or problems with submitted data or metadata. Additionally, other relevant individuals (e.g., names, roles) and bibliographic information (e.g., title, abstract, suggested author list, related publications) make it easier to provide proper attribution about the source of data held in the NODC repository and to provide additional context and aid understanding of the data for future data users.

It is anticipated that the information collected will be disseminated to the public or used to support publicly disseminated information. NODC will retain control over the information and safeguard it from improper access, modification, and destruction, consistent with NOAA standards for confidentiality, privacy, and electronic information. See response to Question 10 of this Supporting Statement for more information on confidentiality and privacy. As a library and archival repository, original data and information archived and stewarded by NODC are not subject to Information Quality Act provisions.

3. Describe whether, and to what extent, the collection of information involves the use of automated, electronic, mechanical, or other technological techniques or other forms of information technology.

The Send2NODC web application requires access to the internet and a web browser to enable a data provider to fill in a series of short form elements by entering free text or by selecting terms

from an NODC-provided list.

The Send2NODC web application is available online for public review at http://www.nodc.noaa.gov/submit/s2n. Because the Send2NODC web application interacts in real time with an NODC-supplied database, it is not effective to print the short forms for offline use. The (expired) NODC Data Documentation Form (paper NOAA Form 24-13) collected essentially the same information as the Send2NODC web application.

The geospatial metadata and data files sent to NODC using the Send2NODC web application will be made available to the public over the internet if those metadata and data files are accepted by NODC for inclusion in the NODC long term repository. Data and metadata are accepted by NODC for inclusion in our repository if those data meet at least minimum documentation requirements and are determined to be within the scope of NODC collection policies (e.g. scientific measurements of ocean water temperature and salinity or photographs of sea birds used to count and identify species for a scientific survey would be within scope, but family vacation photographs taken at the beach would not be within scope).

4. Describe efforts to identify duplication.

The FGDC published a report reviewing eleven ISO 19115 metadata editing tools (see www.fgdc.gov/metadata/iso-metadata-editor-review). All but one of the listed tools require downloading and installing some type of software on a local computer system, often with a monetary or personnel cost for installing, configuring and managing the software. Many of the software packages are from non-US sources and would require significant security reviews to be allowed to be installed on many organizational computer systems. Send2NODC requires no download or installation tasks and is available for use at no cost.

No tools that have an OMB Control Number have been found to exist that provide the two main characteristics of the Send2NODC web application: 1) assisting with creating ISO 19115 compliant metadata to describe a geospatial data set and 2) attaching the described geospatial data files in a bundle that is easily forwarded to the NODC as a request to archive.

5. If the collection of information involves small businesses or other small entities, describe the methods used to minimize burden.

Because Send2NODC is a web application, no software or other technology is necessary other than a standard web browser application (e.g., Firefox, Safari, Chrome) and an internet connection. Send2NODC is designed to make the creation of complete, international standards-compliant (ISO) geospatial metadata simpler for all users by providing drop-down choices for frequently used descriptors (e.g., names of organizations or ships) and eliminating the need for a Send2NODC user to understand the complexities of the ISO 19115 Geospatial Metadata standard. Send2NODC also enables users to attach data files to the metadata created using the web application and send the combined metadata and data to NODC in one web-based Submission Information Package.

6. Describe the consequences to the Federal program or policy activities if the collection is not conducted or is conducted less frequently.

The NODC is widely recognized in the ocean sciences community as the largest archival collection of freely-available ocean observations. Submissions of marine observation data by non-Federal entities or non-Federally-funded entities to be archived at NODC are completely voluntary, but are important for making sure that valuable marine observations are preserved and accessible for future research purposes. Federal entities and Federally-funded entities are expected to make data and metadata, which are collected using Federal funds, accessible and available for the long term in accordance with Executive Order 13642 "Making open and machine readable the new default for government information".

If the information collection supported by S2N for data sent to NODC is not conducted, the result are twofold: 1) oceanographic data, typically collected with significant Federal financial support may not be made discoverable and accessible to a broad audience of data consumers and 2) additional scarce resources will be expended by NODC to create and organize metadata to document each request to archive marine data. The Send2NODC web application is designed to support an 'as needed' or 'on demand' frequency of use, so a data provider may use S2N only when appropriate.

7. Explain any special circumstances that require the collection to be conducted in a manner inconsistent with OMB guidelines.

Not applicable.

8. Provide information on the PRA Federal Register Notice that solicited public comments on the information collection prior to this submission. Summarize the public comments received in response to that notice and describe the actions taken by the agency in response to those comments. Describe the efforts to consult with persons outside the agency to obtain their views on the availability of data, frequency of collection, the clarity of instructions and recordkeeping, disclosure, or reporting format (if any), and on the data elements to be recorded, disclosed, or reported.

A Federal Register Notice published on 15 August 2014 (79 FR 48124) solicited public comments. No comments were received.

Additional comments from the public were solicited during the 2014 American Geophysical Union Ocean Sciences meeting in Honolulu HI. Conference attendees, mostly from the ocean science community, were invited to create a test user identity and test metadata for a Submission Information Package with the Send2NODC web application. NODC staff recorded approximate times needed to create a user account, to complete the forms, and to "send" a test submission package to NODC. Comments about usability features, such as the layout or help guidance on each tab within the application, were recorded and addressed in follow-up design improvements.

In coordination with the Federal Register Notice, NODC provided a voluntary Google-based web form for any additional comments from the public, in addition to comments solicited in the Notice. Of the approximately 90 comments received in the web form, about 50% of the comments were characterized as a recommended (future) enhancement or change, about 25% of

the comments identified an error or problem (most commonly, an error in the messaging related to uploading files to the NODC file staging area), about 15% of the comments were characterized as a question about how the application was working or was supposed to work, about 6% of the comments were characterized as commending NODC on the use or design of the application, and the remaining comments (about 4%) were characterized as 'other' (not fitting any of the other categories). The NODC design team is working through each comment to address the specific errors and problems, to make note of feasibility of recommended future enhancements, and to respond to questions about use or design of the application.

Comments specifically about record keeping requirements, time required for using the application, or the type of information being collected were generally positive statements, e.g., "The time required to make an account was ~2 minutes. The time it took to submit a submission package was ~25 minutes" and "Overall, making a new account was a straightforward process. Took only 5 minutes which is an acceptable length. The required information was appropriate and easy to provide. Also, the instructions and sidebar tips were clear and helpful."

9. Explain any decisions to provide payments or gifts to respondents, other than remuneration of contractors or grantees.

No payments, gifts or other remuneration are planned.

10. Describe any assurance of confidentiality provided to respondents and the basis for assurance in statute, regulation, or agency policy.

No assurances are provided regarding confidentiality. Inclusion of identifying information (e.g., name, email address) in the comments form was voluntary. Identifying information in the Send2NODC web application user profile is required to maintain the account and to comply with mandatory standard metadata requirements for geospatial data in accordance with Executive Order 12906.

11. Provide additional justification for any questions of a sensitive nature, such as sexual behavior and attitudes, religious beliefs, and other matters that are commonly considered private.

Not applicable.

12. Provide an estimate in hours of the burden of the collection of information.

The initial information collection activity is estimated to take an average of one hour, which includes creating a user account, entering at least the minimum information needed to submit a Submission Information Package, reviewing some ancillary information (help text and tips), and submitting the package to NODC. The response to Question 8 describes how NODC arrived at this estimate (e.g., internal NODC alpha tests, external alpha test at professional conference, external beta test announced via Federal Register Notice and NODC home page and social media releases). Comments and observations during these tests indicated that the range of response preparation was from about 0.5 hours (minimal completion) to 1.5 hours (extensive documentation). On average, submissions after the account has been set up will take less than one hour. Also, revisions made to S2N as a result of the beta testing include the ability for a

responder to reuse and modify all information entered during a prior response. This capability is expected to significantly decrease the time requirements for anyone who uses S2N more than one time for similar types of data. However, we will continue to estimate an average of one hour for all responses (when a request for this information collection was last submitted, the burden was estimated at 30 minutes per response).

NODC estimates that 200 distinct respondents will submit at least one response and 100 of those respondents will submit at least two responses, for a total of at least 300 responses. This is approximately the same number of responses anticipated when NODC used the paper-based NOAA Form 24-13 "Data Documentation Form" (however, when a request for this information collection was last submitted, 100 respondents and 200 responses were estimated). A paper-based form and an electronic (Adobe Acrobat PDF) version of NOAA Form 24-13 "Data Documentation Form" are available, but all data providers will be encouraged to use Send2NODC web application and NODC will eventually discontinue use of these representations of NOAA Form 24-13 "Data Documentation Form". The estimate for collection information in these other forms is closer to 30 minutes, as the information is now being requested in a more detailed way.

Graduate student assistants or data managers for senior scientific research staff often create metadata content. Assuming an average labor rate comparable to GS-8 Step 5 ($43,075/year or $20.71/hour in 2014), the estimated labor cost to complete the average response for Send2NODC is $6,213 per year (300 hours x $20.71/hour).

13. Provide an estimate of the total annual cost burden to the respondents or record-keepers resulting from the collection (excluding the value of the burden hours in Question 12 above).

Capital and start-up costs are not applicable because equipment required for this information collection (i.e., a computer connected to the internet, standard browser software, office management software for recording information, etc.) are standard office/laboratory equipment managed and operated by all anticipated respondents. Operations and maintenance costs for required equipment (listed above) are not applicable because these costs are part of routine business operations. All information collection transactions are conducted through the web application interface or via email, so that additional costs than those already paid for internet connectivity will be incurred.

14. Provide estimates of annualized cost to the Federal government.

Collecting, reviewing and updating metadata about scientific data sent to NODC is an essential function of the archival process. When necessary, NODC staff (Federal and non-Federal) may create metadata sufficient to document data sent to NODC to be archived for the long term.

Current testing indicates it will take NODC staff approximately one hour (on average) to review and begin using metadata and data files that are the response to this information collection through Send2NODC web application. This does not include additional tasks that are required for completing archival ingest processes that are not part of the Send2NODC web application information collection activity.

Assuming an average labor rate comparable to GS-11 Step 5 ($71,504/year or $34.38/hour in 2014 in the Washington DC metro area), the estimated labor cost to the Federal government for processing the average response through Send2NODC is $10,313 per year (300 hours x $34.38/hour). If Send2NODC is not used to collect information from a data provider, including uploading/including data files through the Send2NODC interface, the estimated labor time increases to at least 1.5 hours, for an approximate annualized cost of $15,470/year.

15. Explain the reasons for any program changes or adjustments.

The overall response time is now estimated at 1 hour. Previously, with the simpler paper form, the time was estimated at 30 minutes. Because the majority of researchers are submitting their information via Send2NODC, which allows a more detailed response, we are now assuming 1 hour per response. Since there are both more responses and an increase in estimated time, we are considering 100 of the added hours to be an adjustment, and 100, a program change.

16. For collections whose results will be published, outline the plans for tabulation and publication.

Metadata and data sent to NODC in response to this information collection via Send2NODC web application are typically discoverable by the public within 3 days of receipt of the submission information package through the NODC web site, geoportal and other web discovery services. Additional access by the public to metadata and data through web catalogs and other portals (e.g., data.gov) becomes available upon completion of additional archival ingest processing, which is not directly related to the Send2NODC web application.

17. If seeking approval to not display the expiration date for OMB approval of the information collection, explain the reasons why display would be inappropriate.

Not applicable.

18. Explain each exception to the certification statement.

Not applicable.

B. COLLECTIONS OF INFORMATION EMPLOYING STATISTICAL METHODS

The Send2NODC web application does not employ statistical methods.