

Task Order # L-13

Climate Database Modernization Program

July 1, 2003

TASK ORDER TITLE:

NARA Imaging Tasks- Lightships and Meteorological Journals

1.0 Introduction

This project involves two separate tasks to be completed in two different phases. The first task will involve scanning the historical lightship logbooks which will be shipped in increments from Archives I to Archives II during the course of this task order. The additional task will be to scan various historical meteorological journals and records when the lightship scanning project is complete. Lason staff involved in this project, will attend a NARA training course on handling fragile manuscript records and documents prior to beginning the imaging.

2.0 Phase 1. -Lightship logbooks

This project involves scanning mainly Atlantic lightship logbooks from Archives I, although, the collection will include some from the Pacific Ocean and Great Lakes. Where possible, the order of scanning will follow the priority list produced by K. Elder of Woods Hole Oceanographic Institute (WHOI) (Attachment 1 of this task order) which gives Atlantic lightship records with the longest station histories the highest priority. The total collection is in 435 boxes and all boxes/logbooks will be imaged under Phase 1.

During mid July 2003, thirty to fifty boxes of lightship logbooks will be shipped to Archives II. The entire series of Lightship Records (435 boxes) in box number order will be shipped to Archives II (College Park) in three shipments. NARA expects that all of the boxes will be at Archives II by mid-August, perhaps sooner. The Lason people will begin scanning the material from the first shipment on July 16 using the excel inventory (attachment 1) provided. Lason shall provide a monthly summary in their progress report of what boxes/stations and total number of images that have been scanned.

Of the 435 boxes, the finding aid identifies 273 of them as containing east coast (Atlantic) lightship records. The excel spreadsheet list contains the known Atlantic lightships and their WBAN numbers, however, the inventory gives only the station name of the first and last log book per box and those between have no index. For this reason, there could be Atlantic logs between those boxes indexed as containing West Coast or Great Lake stations that do not appear in our inventory/priority list. Additional information on assigning WBAN numbers for the non-

listed ones is given under the daily/monthly index structure portion of this document. Non-east coast lightship records will also be imaged and indexed in detail and added to the excel spreadsheet database.

2.1 Lightship logbook -Index Structure- Daily and Monthly Formats

We believe the majority of the forms will contain daily data, however there will be monthly summary forms that will also need imaging. These daily and monthly forms will be archived in two separate cabinets on WSSRD because of the date difference resulting in slightly different indexing structures. The daily and monthly image files must be maintained separately for loading to WSSRD.

Daily Index Structure

Batch number - Assigned by contractor to the lightship stations, incorporating the NARA Box number inventory. The batch numbers shall incorporate the NARA box numbers (i.e. NARA 001 - NARA 435 is the range of batch numbers to use) as part of the 7 character batch ID for each logbook within that box. It is possible for the same batch number to appear in both the daily and monthly cabinets.

WBAN Number -See station list, a WBAN number has been assigned for each known Atlantic lightship station and is included in the excel spreadsheet (Attachment 2). Insure the dates match for the period when the WBAN number was effective. For lightship stations that don't have a WBAN number assigned, please assign one starting with a negative number in numerical order (i.e. -0001, -0002, etc.). A list of station names and assigned negative WBAN numbers is to be provided to NCDC upon completion of the logbook imaging at NARA. Contact NCDC at 828-271-4499 or tom.ross@noaa.gov if any conflicts are noted.

Vessel Number- This may or may not appear on each page with the data. If unavailable leave this field blank.

Vessel Name - This generally appears on each page with the data, but it is often difficult to read clearly; you can check the name against the excel spreadsheet for the correct spelling of the Atlantic lightships listed. For those not on the Atlantic list, once the name has been established and a WBAN number assigned insure that a standard entry is maintained in the indexes.

Date - The date is to be entered as MM/DD/YYYY in this index field. Each image that contains data generally has the date entered in the upper right corner. This is critical information and if not available on the page should be determined if possible by examining the logbook pages on either side.

Form Type - Usually located at the top of the data page, if not available, enter "Unknown"..

Load Date and Time- Internal to WSSRD

Document ID- Internal to WSSRD

Monthly Index Structure

Batch number - Assigned by contractor to the lightship stations, incorporating the NARA Box number inventory. The batch numbers shall incorporate the NARA box numbers (i.e. NARA 001 - NARA 435 is the range of batch numbers to use) as part of the 7 character batch ID for each logbook within that box. It is possible for the same batch number to appear in both the daily and monthly cabinets.

WBAN Number -See station list, a WBAN number has been assigned for each known Atlantic lightship station and is included in the excel spreadsheet. Insure the dates match for the period when the WBAN number was effective. For lightship stations that don't have a WBAN number assigned, please assign one starting with a negative number in numerical order (i.e. -0001, -0002, etc.). A list of station names and assigned negative WBAN numbers is to be provided to NCDC upon completion of the logbook imaging at NARA. Contact NCDC at 828-271-4499 or tom.ross@noaa.gov if any conflicts are noted.

Vessel Number- This may or may not appear on each page with the data. If unavailable leave this field blank.

Vessel Name - This generally appears on each page with the data, but it is often difficult to read clearly; you can check the name against the excel spreadsheet for the correct spelling of the Atlantic lightships. For those not in the Atlantic, once the name has been established and a WBAN number assigned insure that a standard entry is maintained in the indexes.

Date - The date is to be entered as MM/YYYY in this index field. Each image that contains data generally has the date entered in the upper right corner. This is critical information and if not available on the page should be determined if possible by examining the logbook pages on either side.

Form Type - Usually located at the top of the data page, if not available, enter "Unknown".

Load Date and Time- Internal to WSSRD

Document ID- Internal to WSSRD

3.0 Delivery Media

The images and indexes are to be provided to NCDC on CD-ROM, which will be forwarded to IMC for loading to WSSRD. A backup copy of all the delivered files (images and indexes) must be maintained by Lason until the project is officially completed. NCDC will then notify Lason as to the disposition of these backup files. A batch shall never be split between two separate CD-ROMs. Similarly, the two collections (daily and monthly) shall always be maintained separately, never being mixed together on a single CD for delivery.

The CD-ROM Image path shall follow the format: "P2000\NARA001\12345678.tif"

(Note: basically the same procedures as used for the serial publications)

Where:

A. P2000 represents the volume name for the 1st CD-ROM. Each additional CD-ROM volume shall be increased by one. Volume numbers P2000 through P2999 have been allocated for this task;

B. NARA001 is the batch number to be assigned as described above under the indexing structure.

C. 12345678.tif is a unique image filename to be assigned by the contractor.

4.0 Quality Assurance

The images and indexes must conform to the standards laid out in the CDMP "Imaging, Indexing, Keying & Storage Standards" dated April 30, 2002. Once loaded to WSSRD the images and indexes will be reviewed by NCDC and (K. Elder-WHOI) using the WSSRD Quality Assurance Module. Any failed batches must be corrected at the batch level for re-loading to WSSRD.

5.0 Logbook Return

The scanning work on this project is being accomplished on site at Archives II. All imaging of the lightship records and all documents to be imaged under phase 1 and 2 shall be completed by the end of the contract period (March 31, 2004).

6.0 Phase 2. -Smithsonian Meteorological records and Journals

When the scanning of the lightship records are complete, Lason will move to Phase 2 of the project which will involve scanning Smithsonian and additional miscellaneous records and journals. These volumes have been basically identified, although a few additions may be added before the beginning of Phase 2. NCDC will provide a copy of this file (priority list of

volumes) prior to the beginning of phase 2. The delivery date shall be coordinated between Lason and NCDC.

Lason shall contract, if an agreement can be reached, with NARA microfilming staff to film the Smithsonian records (9 oversize boxes). Lason shall then image the records off the microfilm in their Beltsville office. These will likely be indexed the same as the forts documents.

Lason shall then concentrate on imaging the remaining documents on the priority list using their Zeutschel book scanner. Generally, each journal will be indexed in a similar manner as the Ludlum collection of journals and diaries. This format includes: Station name, State, Observer, Source, Beginning Date, and Ending Date. The majority of records will probably be multi-page tif files. However, some records may be indexed down to the month or even daily level if they contain a long time series of daily data.

Lason shall provide NCDC sample copies of all the unique records imaged while in NARA from the priority list on CD-ROM for review prior to indexing so the index structures can be established. Without good knowledge of each document or journal NCDC can not develop a sound indexing structure. The samples for each collection (e.g. the Smithsonian) can be provided upon completion so the indexing can begin before all the imaging (all collections) is completed at NARA. If the NARA document imaged is a multi-page publication or journal the sample images should include the title page, table of contents, and several representative pages from the volume. NCDC will work with Lason to establish a schedule to reduce any loss in indexing production. Also if the contractor has any suggestions on how NCDC might better view the imaged material and develop the indexes NCDC would most welcome the input. This paragraph is only a suggested approach on how the project should proceed and before a final decision is formulated both NCDC and the contractor should discuss possible methodologies for developing the index structures.

The delivery media and quality assurance will be same as for Phase 1.