THE NORTH ATLANTIC RIGHT WHALE CONSORTIUM
Data Sharing and Use Protocols and Conditions
Updated 2025-04-25 with revisions in red text

INTRODUCTION

The North Atlantic Right Whale Consortium Databases were established in 1986 as part of a cooperative right whale research program conducted by the University of Rhode Island, New England Aquarium, Center for Coastal Studies, Woods Hole Oceanographic Institution, and other organizations forming the North Atlantic Right Whale Consortium. The data are comprised of two major datasets, the survey and "Sightings Database," maintained and curated by the University of Rhode Island, and the "Identification Database" (formally the "Photo-Identification" Database), maintained and curated by the New England Aquarium, as well as several other smaller databases. Records contained within these databases come from a wide variety of contributors that includes research groups, management agencies, whale watchers, and individual mariners.

The **Sightings Database** contains records of thousands of sightings of right whales in the North Atlantic Ocean, as well as sightings of many other species of whales, dolphins, sea turtles, seals, and large fishes. It also contains survey data associated with many (but not all) of these sightings that allow quantification of associated survey effort (**note that the database does not include interpreted effort data as such**). Though most sightings in the Sightings database are from surveys conducted from the late 1970s to the present, some right whale historical records go back as far as the 18th Century. Each record in the Sightings Database represents a group of animals (i.e., a group of 3 whales has a single record just as a group of 1 does) and there may or may not be photographic proof of a given sighting, including right whales.

The **Identification Database**, also referred to as the North Atlantic Right Whale Catalog, or "Catalog" for short, contains all photographed sightings of right whales since 1935. In addition to photographed sightings, the database now contains any record that can lead to an individual identification. This includes "sightings" with just a skin sample and no photographs since that whale can be individually identified by genotyping. It also includes one quality location per day from satellite tags attached to identified individuals. Photographed sightings are matched to whales in the Identification Database whenever possible so that individual animals can be monitored over time. Each record in the Identification Database is of a single individual (i.e., a group of 3 whales will have 3 separate records with an association code linking the 3) and each record contains time, date, location, observer, notes and behaviors.

The Sightings and Identification Databases are periodically cross-referenced, so that individual identification data from the latter can be linked to sighting data from the former. For that reason, many sightings in the Identification Database are eventually included as opportunistic sightings in the Sightings Database (with an approximate 1-year lag). Although the animals' identifications are not included in the Sightings Database, the two databases can be linked on common fields.

WhaleMap is an open-source software system that was designed to collate and disseminate the latest right whale observations and survey results in near real-time to inform more effective, dynamic planning of research and conservation activities. It ingests data from the repositories of a number of different survey groups, converts the data to common format, and makes it available for online display so that shortly after their planes land, boats tie up, or autonomous vehicles call home, the survey results will publicly available. The online display, as well as additional information on the system, is available at https://whalemap.org. These are preliminary data

that are subject to change. As a result, access to these data is limited and requires clear justification as to why they are required rather than the quality-controlled data in the Identification and/or Sightings Databases.

***Please Note**:*
Displays of whale sighting locations do not represent where whales are present; the locations are obsolete within minutes because whales swim continuously. Sighting locations are also limited to where an observer was present when a whale came to the surface. This coincidence is exceedingly rare because observers are present in only a very small fraction of whale habitat for very limited periods of time. And even dedicated marine mammal surveys conducted in ideal weather conditions are likely to miss whales because whales spend most of their time out of sight below the ocean surface.

Statistical models can provide a more informative map of whale presence in an area by taking into account observer effort, weather conditions, and whale dive times. Models can also evaluate correlations between whale sighting locations and environmental factors, and use identified relationships to predict whale presence in areas or during times with no observer effort. NOAA uses this Duke model of right whale seasonal densities along the US East Coast to develop plans to reduce entanglement and vessel collision risks.

Lastly, there are several **other databases** held at various institutions that may be accessed via NARWC request, including: anthropogenic events, genetics, blubber thickness measurements, contaminant levels, visual health assessments, and necropsy findings. These data are all linked to the Identification Database by animal identification number and/or sighting information from the Identification Database. Please note - Requests for biological samples may require additional permitting obligations that must be met before samples are accessed. Since the materials come from numerous independent individuals and institutions, these databases are not strictly proprietary. Rather, they represent a scientific resource, and access to the data for scientific, educational, conservation and management purposes is encouraged. Contributors have full and unrestricted access to use of their own data as well as to the life-history data of the individual whales they documented. Proposals for data access from scientists, managers, students or other individuals with a bona fide purpose will be reviewed by the Consortium Board members and/or Database Curators. Given the great effort required to collect the available data, the Consortium and the curators of the data have an obligation to protect the rights of contributors by placing certain restrictions or conditions upon access to, and use of, the materials within it. For this reason, access to data, both historical and future, should not be assumed. Those wishing to access NARWC data should make all reasonable efforts to do the following **in advance of project development**, including, but not limited to, funding proposals, postdoctoral research, student projects, and permit applications:

1. Identify major data contributors of desired dataset(s)
2. Connect with those data contributors to determine data availability and/or potential capacity for collaboration

NARWC administrators and data curators (contact information is included at the end of this document) can assist with identifying appropriate data contributors. While connecting with **all** data contributors may not be possible, particularly for range wide/broad scale analyses, efforts should be made to connect with those contributing significant portions of the desired dataset(s). This is particularly important for regional assessments/analyses.

These restrictions/conditions are not intended to be onerous, rather, to ensure that data requesters understand nuances of the data, that data are used and interpreted correctly, and that we avoid duplication of efforts that may interfere with the publication of existing work. The Consortium Board and Database Curators are

committed to working with requesters to develop and modify requests in order to honor the founding Consortium principals of data sharing and collaboration.

DATA ACCESS PROTOCOLS

Data access may be requested from scientists, managers, students, or other individuals with a *bona fide* purpose. Data access will not be granted for open-ended, exploratory investigations, or for any use that would make data freely accessible to the public.

In order to ensure that research being planned or currently conducted by contributors is not compromised or unnecessarily duplicated, and that proper authorship or acknowledgment of all major data contributors occurs, any request for data must be submitted to the Consortium via the online request form located on the NARWC website. The proposal itself need not be lengthy, but it should at a minimum contain sufficient information on the following:

- Name of the requesting institution(s) and of the Principal Investigator;
- Outline of the proposed work, including working title, questions being addressed, hypotheses tested or anticipated management application;
- Anticipated data requirements (and their relevance to the proposed work) including specific data fields requested and both the time period and geographic scope of interest;
- Identification of major contributors to the requested dataset and confirmation of communication with said contributors
- Justification for the data requirements in the context of the project and if requesting WhaleMap data, a justification of the need for unprocessed data;
- Anticipated products of the work (e.g., scientific paper, student thesis, environmental assessment (EA), environmental impact statement (EIS), management plans, reports, derived datasets);
- Estimated time frame to completion of the study (not to exceed two years). A report summarizing the work will be due 6 months following the project end date. It is understood that peer-reviewed and thesis publications may require a longer time frame and so a progress report indicating the outcome of the study may be filed while awaiting publication).

General review procedure for proposals for publication purposes:

Publication type requests include peer-reviewed publication, thesis work, popular news, educational/curriculum projects, and others. Publication proposals are typically reviewed within thirty days of submission by Consortium Board members with knowledge of the type of work being proposed and/or a curatorial role with the data. Please be aware, however, the reviews may take longer depending on the complexity and size of the request. Additionally, the need for clarification on the scope, data requirements, or other aspects of the request may delay the review process. In some cases, proposals will also be sent for review to those organizations that contributed substantial portions of the data being requested. The review will be focused on ensuring that duplication of effort is minimized, that proposed analyses seem appropriate, and that potential coauthors are identified. Reviewers will review the proposal and discuss appropriate authorship given their knowledge of who contributed the majority of the data required for the proposed project. Their recommendations for authorship will be sent to the applicant via the Consortium Executive Administrator. In some cases, the reviewers may suggest that, instead of authorship, acknowledgement of the Right Whale Consortium as a whole and/or certain institutions/persons be included in any published document. The applicant must ask the proposed author(s) whether they would be interested in authorship and let the Consortium Executive Administrator know what their responses are. Once authorship has been agreed upon (among the applicant, the authors, the executive

administrator and the reviewers), the data will be released. If there is a disagreement about authorship, it will be resolved by the board. The applicant will then be provided with the requested data, with the method (email, cloud-based server) determined by the size of the requested information file. Although the curatorial workload is heavy, and researchers are often very busy, every effort will be made to provide information in a timely fashion. In most cases, there will also need to be some follow-up discussion between the applicant and the data curator on the exact details of the data required and specific formats. For data extractions that require novel extraction code or additional database programming/analyses, there may be associated fees. Major data contributors will be notified that their data are included in an approved request and may have additional insights for the access requesters regarding data nuances.

Requests for educational purposes within the context of class projects, but excluding thesis, seminar, research symposia, and/or capstone projects, may be expedited through the use of standard datasets. This determination will be made by the NARWC admin and data curators. In addition to the data access conditions outlined below, products resulting from these requests may not be disseminated outside of the class for which the work was done, including, but not limited to, social outlets.

The curator and the reviewers will treat proposals as confidential and ideas or hypotheses that they may contain will not be shared with third parties. The only exception would be if the Consortium reviewers wish to obtain confidential peer review of the proposed work in order to judge its feasibility or merit; this would only be done with prior approval of the applicant. Conflicts over the use of the data will be mediated by Consortium board members in as timely a fashion as possible. The Consortium will encourage multi-investigator proposals where interests of several investigators overlap.

Grounds for the rejection of a proposal will include lack of investigator qualifications, lack of specificity of data requirements, requests for wholesale database downloads, assessment that the scope of the project is unreasonably large or not feasible within the proposed time frame, unwillingness of the investigator(s) to acknowledge or offer authorship to major data contributors, a determination that the proposed work is already underway by someone else, and/or if the applicant has violated a past data access agreement.

General review procedure for proposals for management purposes:

Proposals for management purposes only will be transmitted by the Executive Administrator to the appropriate data curator(s). Management proposals may include environmental assessments (EA), environmental impact statement (EIS), management plans, and other similar type requests. In general, management requests are not meant for broad public distribution, though the final report/product (not the data) may be publicly available. In general, management requests are typically reviewed by the database curator(s) and not by full Board review, however, a review may be requested based on the extent and final product plan. The applicant will be provided with the requested data, with the method (email, cloud-based server) determined by the size of the requested information file. Although the curatorial workload is heavy, and researchers are often very busy, every effort will be made to provide information in a timely fashion. In most cases, there will also need to be some follow-up discussion between the applicant and the data curator on the exact details of the data required and specific formats. For data extractions that require novel extraction code or additional database programming/analyses, there may be associated fees. Major data contributors will be notified that their data are included in an approved request and may have additional insights for the access requesters regarding data nuances.

Additional review protocols for WhaleMap data:

Access requests for WhaleMap data will also be reviewed in consultation with the contributing data

organizations. Any data that are incorporated into WhaleMap from other platforms, including Canada's Whale Insight, will follow additional protocols as requested by those sources.

DATA ACCESS CONDITIONS

Provision of any data will be made subject to the conditions given below, to which the applicant must agree within their proposal. These conditions are designed to eliminate misunderstandings, and to protect both the applicant and the organizations that curate the data, as well as the data contributors. Conditions vary slightly for management use, so careful attention should be given to specific conditions.

Users of Consortium data are encouraged to notify specific database curators of data quality issues and/or edits.

To the extent practical, databases that result from the proposed studies, that are appropriate for archiving and subsequent use within the consortium data access process, can be curated by the Executive Administrator. Subsequent use by others would be at the investigators' agreement.

Conditions for data use intended for publication (includes peer-reviewed publications, thesis work, education modules, and other uses that are not strictly managerial):

- For a reasonable period of time (generally that of the estimated time frame of the applicant's proposed study, not to exceed two years), the Consortium will not provide similar data to others for the same or similar scientific purposes described in the applicant's proposal, without first obtaining the applicant's permission. Exceptions to this may be made if the access request is for substantial datasets and/or efforts are similar in nature but utilizing alternate approaches. This exception will often apply to modeling and AI efforts.
- The applicant will use the requested materials for only those purposes set forth in their proposal. Requests for significant departures from the scope of the proposal must be submitted via the NARWC data access form.
- The applicant will not share the requested materials with any third party.
- For journal publication data repository requirements, the applicant will work with the NARWC Executive Administrator and data curators to determine that most appropriate level of aggregated data to supply. Additionally, the NARWC database curators maintain copies of data files shared for access requests.
- The applicant agrees to complete the work in the time frame given, although requests for reasonable extensions of this time frame will of course be considered.
- The applicant agrees to publish the results in a refereed journal in a timely manner. Given the endangered status of the North Atlantic right whale population, it is critical that any research results that provide a better understanding of the biology and behavior of this species be disseminated so that they are accessible to other scientists and managers. Failure to do this constitutes unfair monopolization of data with no benefit to the population. **A written report must be submitted to the Consortium Executive electronically within six months of the specified end date of proposed work. Manuscripts may serve as the report.** Failure to supply a report would preclude further data access. Reports will be accessible by Consortium Board Members. Additionally, researchers are strongly encouraged to present right whale related work at the Annual North Atlantic Right Whale Consortium Meeting.

Conditions for data use intended for management purposes only:

- Applicants may use the data for other management-related analyses on one condition: they inform the Consortium Executive Administrator of **each additional project**. This process allows the Consortium to establish links between the applicant and other managers and/or scientists interested in similar analyses and

also permits the data curator to inform the applicant about available updates to the data or errors discovered in the data subsequent to its provision. Also, by tracking the different ways the data are used, the Consortium can further illustrate the benefits of shared data. Although persons other than the initial applicant may perform the additional analyses, it remains the responsibility of the initial applicant to inform the Consortium of the additional work.

- In all cases where the NARWC database is used for management purposes, the following disclaimer must be included in the document: "**Raw sighting data from the NARWC database are not effort-corrected and the management documents in which they are used are not peer reviewed. Distributional patterns based on these data are likely to be biased by where, and when, surveys were conducted.**" The data curator will inform the applicant of any limitations of the data provided (e.g., from an area with little or no survey effort, for a species that has low detectability from the surveys available, from surveys with strong seasonality, or with other known biases).
- In cases where the document (not peer-reviewed) presents any additional interpretation based on sightings data made available by the Consortium, the producer of the report must identify the relevant discussion and interpretation sections referring to the database in a transmittal letter prior to public release. In such cases, one or more members of the Consortium Board may be tasked to review these sections. These reviews will be limited to ensuring that the data are not misused or misrepresented. In such cases where reviewers feel that data have been misused or misrepresented, the Consortium will submit a public comment, limited to the use and interpretation of the Consortium data, to ensure that an accurate correction is a part of any official public record and the decision-making process.
- If the management analyses result in publishable information, the applicant is required to submit an additional request for publication. If someone has already applied for data to publish on a similar analysis, the Consortium will encourage a dialog among the parties, but publication rights will go to the applicant who first applied for data under the publication request process.
- **A written report must be submitted to the Consortium Executive Administrator electronically within six months of the specified end date of proposed work and may include** any document or other product produced using the Consortium data provided including but not limited to EIS's, EA's, Take Reduction Team documents, PowerPoint presentations, Endangered Species Act Section 7 consultations, workshop proceedings, executive summaries, or other unpublished "gray literature." Reports will be accessible by Consortium Board Members.

All users of NARWC data must acknowledge and agree to the condition of data access outline above.

Preferred Citation of Consortium Data

Each database accessed should be given an individual citation. Citations of Consortium data should follow the format below:

In text:
North Atlantic Right Whale Consortium (YEAR)

In Literature Cited:
North Atlantic Right Whale Consortium (YEAR). North Atlantic Right Whale Consortium XXX Database MM/DD/YYYY (Anderson Cabot Center for Ocean Life at the New England Aquarium, Boston, MA, U.S.A.).

XXX could be: Identification, Sightings, WhaleMap, Genetics, Contaminants, Necropsy, etc.

If you have any questions, please contact the North Atlantic Right Whale Consortium Executive Administrator at *narwc@neaq.org*.

Database Curators

Contact information for database curators is hyperlinked below.

Sightings Database
Identification Database
Acoustics (coming soon)
Others (descriptions located in the access protocols)
Anthropogenic Events (Entanglement, Vessel Strike)
Biopsy Archive
Blubber Measurements
Contaminants
Genetics
Necropsy
WhaleMap
Visual Health