

MarineGEO Data Policy & Authorship Agreement

v 0.1.0

Summary

- Most data are made available to all participants in the network immediately after basic quality control.
- Data are typically made public within 2 years of submission.
- Use of data prior to its public release requires pre-approval by the MarineGEO Central Team and proper attribution on any scientific products.
- Participants who make substantial contributions to MarineGEO activities will be eligible for co-authorship on resulting publications and products.

Overview

MarineGEO values collaboration and open sharing of our collective knowledge. Acquiring, curating, sharing, and analyzing data are central functions of our network's research. This document summarizes guidelines for data contribution, sharing, and authorship by contributors to activities coordinated by the Smithsonian-led MarineGEO program. Adherence to the data policy described here is a requirement for any use of MarineGEO data or participation in any MarineGEO activity that involves data collection.

Terms & Definitions

MarineGEO refers to both the Marine Global Earth Observatory research program and the broader community (network) of partners that conduct its activities, which are led and coordinated by the MarineGEO Central Team at the Smithsonian Institution.

MarineGEO Central refers to the leadership and support team, based at the Smithsonian, that coordinates MarineGEO activities. MarineGEO Central is responsible for collating, curating, archiving, and serving MarineGEO data to contributing participants and, where appropriate, to the public.

Data refers to all content generated by MarineGEO activities, including measurements stored in spreadsheets and on original field sheets, images, video, genetic sequences, GIS layers, computer code, and associated metadata.

A *contributing participant* is any party who has collected and contributed original data to MarineGEO. Participants can include principal investigators, technicians, students, and/or volunteers at MarineGEO partner institutions, as well as those contributing to a MarineGEO project by mutual agreement.

Data Collection

MarineGEO provides standard protocols and data formats for most network activities. Contributing participants are responsible for ensuring the data they collect adheres to MarineGEO best practices. Contributing participants agree to utilize these resources provided by MarineGEO, including protocols, field sheets, and data entry spreadsheets.

Data Submission

Contributing participants are responsible for using MarineGEO's data submission guidelines and portal to submit their data for curation in a timely fashion (usually within six months of collection unless otherwise noted). Contributors also are responsible for responding to any requests that arise during the Quality Assurance--Quality Control (QAQC) process to ensure data quality and completeness.

Data Availability

MarineGEO strives toward open-access data. However, we recognize there are different categories of data (Table 1) and contributors deserve priority in their capacity to utilize certain types of data they contribute. In general, data collected as part of MarineGEO activities will be made available to contributing participants immediately upon meeting all requirements for data quality. In addition, participants will have access to all data relevant to the project(s) in which they participate. Certain types of data, including products of experiments and student thesis research, can be embargoed for a term specified by mutual agreement before public release.

Metadata associated with MarineGEO data will be made available to the public along with the data, except for certain types of sensitive data such as locations of vulnerable species or habitats, and Personally Identifiable Information (PII), including names, contact information, and images of people. *If you do not wish for such information to be made public, please indicate so during data submission.*

Approximately two years after the initial submission, MarineGEO intends to make all applicable data available on public repositories with proper attribution (see Table 1 and Data Attribution).

Data Sharing

Contributing participants own the raw data they collect.

MarineGEO may analyze, synthesize, and share data as needed to advance the MarineGEO community's collective goals, provided this does not conflict with other guidelines described here. Data may be shared with external parties prior to public release to assist in data curation, analysis,

and/or synthesis as needed to ensure quality and accessibility at the discretion of MarineGEO Central Team.

Sharing data collected by another MarineGEO partner with parties outside the network prior to public release requires the prior written approval of the contributing participant(s) who generated the data and the MarineGEO Central Team.

Data Use

Contributing participants are free to use any MarineGEO data to which they have access for scientific purposes (e.g., manuscripts, dissertations, theses) within the guidelines described in this document, or unless expressly restricted by agreement for a particular use.

Where data are collected for a particular use (e.g., a collaborative MarineGEO project), those data may not be shared outside the network or published until the collaborative project is concluded, unless the project participants and MarineGEO Central Team grant prior approval.

Data may not be used for grant or contract proposals, or for commercial or for-profit applications, without the prior written approval of MarineGEO.

Data Attribution and Publication

All products developed using data collected for and/or provided by MarineGEO must provide proper attribution in the form of a written acknowledgment citing MarineGEO and any contributing partners as the source of the data. Where available, these products must also cite the data set's Digital Object Identifier (DOI).

Authors of peer-reviewed publications and other products using MarineGEO data must request and register a MarineGEO contribution number prior to publication and cite it in the acknowledgements of the publication (see: <https://marinegeo.si.edu/publication-request-form>).

Data Updates

Data may be updated periodically to fix errors, update nomenclature, and adhere to current standards. MarineGEO will strive to make the identity of the latest version and the history of all data sets clear. It is the user's responsibility to ensure that they are working with the most current version.

Data Liability

All data provided by MarineGEO are made available "as is." It is the sole responsibility of the user to ensure that the data are being used and interpreted properly, and that they adhere to the terms of this agreement. Neither the contributing participant(s) nor the Smithsonian Institution are liable for any damages resulting from any use or interpretation of the data.

Authorship

The opportunity for co-authorship on publications and products resulting from MarineGEO activities will be extended to contributing participants according to typical norms of scientific societies and journals. Generally, co-authors shall meet both of the following qualifications:

- (1) Submit quality data that is used in the publication; and
- (2) Provide substantial input to lead author(s) during preparation of the manuscript.

Contributing participants who provide additional services, such as submitting a relatively large amount of data, conducting key analyses, and/or contributing disproportionately to writing and manuscript preparation, may be recognized with priority ordering in the author list. Otherwise, co-authors should be listed alphabetically.

It is the joint responsibility of the lead author of a manuscript and the Principal Investigator of each contributing group to ensure all individuals who meet criterion #1 are identified early in manuscript preparation and offered the opportunity to meet criterion #2 for co-authorship.

DRAFT

Table 1. Description of data types and their availability to contributing participants (the network) and their public release.

Data category	Data type	Details	Availability	
			To Participants	To Public
Geophysical	Maps	Drone imagery, GIS layers	Immediate after basic QAQC	Immediate after basic QAQC
Environmental	Geophysical	Sonde, meteorological station	Immediate after basic QAQC	Immediate after basic QAQC
Biodiversity	BioBlitz	Species names, images, DNA barcode sequences, associated metadata	Immediate after basic QAQC	2 years
Biodiversity	Bio-monitoring	Surveys (e.g., quantitative abundance by taxon)	Immediate after basic QAQC	2 years
Biodiversity	Experiments	Coordinated experiments	Immediate after basic QAQC	At publication*
Personal Information	Personally Identifiable Information	Names, contact information, photographs, and other personal details	Only with signed permission of person [†]	Only with signed permission of person [†]
Metadata	Metadata	Descriptions and summaries of data sets	Immediate after basic QAQC	2 years

* No use in individual manuscripts before network manuscript is accepted, unless approved by MarineGEO Central Team

[†] Smithsonian Institution (SI) requires that non-SI individuals sign a waiver before appearing in images displayed by SI to the public