Society of Vertebrate Paleontology
9650 Rockville Pike • Bethesda, MD20814-3998
Phone:(301) 634-7814 • Fax:(301) 634-7455
Email: svp@vertpaleo.org• Web:www.vertpaleo.org
FEIN: 06-0906643

Best Practice Guidelines for Repositing and Disseminating Contextual Data Associated with Vertebrate Fossils

(First edition released December 2016; Revised December 2019)

I. Introduction

Reproducibility is an essential part of all science, including paleontology. The Society of Vertebrate Paleontology (**SVP**) strives to promote access to scientifically important vertebrate fossils and their contextual data to this end by permanently placing them in public trust repositories. This document describes professional best practices for repositing and disseminating contextual data associated with vertebrate fossils, especially locality and stratigraphic information. A working group led by SVP's Government Affairs Committee assembled the original version of this guide with input from field paleontologists, museum researchers, collection managers, and database managers. This document, first released on December 27, 2016, and subsequently revised once (this version), is drafted with the goal for researchers and repository personnel to facilitate preservation of paleontological resources in the field and maximize the scientific value of research collections now and in the future. These guidelines draw on recommendations from the Global Biodiversity Information Facility (GBIF) *Guide to Best Practices for Generalising Sensitive Species Occurrence Data* (http://www.gbif.org/resource/80512). Although the best practices presented here are meant as guidelines especially because they may not necessarily apply to all possible circumstances, SVP expects members to be aware of these universal professional standards as they develop and to provide the Society with feedback to help improve them.

II. Principles

As required by SVP's Code of Ethics (Article 12, SVP Bylaws), pertinent contextual data, including location and stratigraphic information, should be accurately recorded for every vertebrate fossil at the time of collection to augment its scientific value. These data should be available to allow recollection of the exact sample point, and published where possible as well as being archived along with the specimens in an appropriate, publicly accessible and permanent repository (Article 12, Sections 1 and 4 of Bylaws). In cases where release of these data will not create risk of harm to, or theft or destruction of, resources remaining at the collecting site (sometimes including non-paleontological resources), they should be disseminated freely to facilitate research, education, resource management, and other public benefit uses.

A. Contextual data must be collected to the fullest extent possible when a specimen recovered, including both geographic and stratigraphic information.

B. Contextual data must be archived with each specimen in a public-trust repository whose mission is to make both specimens and contextual data permanently accessible for scientific research, education, and other uses. Fossil locality data should be recorded to the greatest accuracy possible and fully georeferenced, down to meters or even centimeters where possible, and should be reposited along with stratigraphic and other metadata, detailed maps, measured sections, field notes, photographs, and other associated documents. The data should be precise enough that future researchers will be able to return to the collecting site when additional information or specimens are required. For fossils collected from regulated properties (e.g., U.S. Federal lands), the contextual data should also be reported to, and archived with, each relevant land management agency.

C. Research publications that use contextual data should clearly identify the repository that houses the associated specimens and should explicitly reference the corresponding specimen and site catalog numbers so that interested parties can locate the original specimens and their associated data (i.e., to provide reproducibility of published scientific results).

D. Wherever possible, contextual data stored in repositories, including unpublished forms, should be disseminated freely and widely.

E. In some cases, public access to contextual data, especially the precise location of the collecting site, can result in harm to fossils, contextual information (e.g., taphonomic or sedimentologic data), on-going research, or to non-paleontological resources (e.g., endangered species or delicate ecosystems) that remain in the field. In such cases, distribution of information may need to be controlled although the presumption remains in favor of release. Any restrictions placed on the dissemination of contextual data should be adhered to rigorously by the collector, repository, and all parties with whom the data have been shared.

F. The sensitivity of all contextual data, especially the location of the collecting site (criteria defined below), should be reviewed by the permitter (i.e., governing body responsible for, or the owner of, the land where the fossils were collected), the repository, and the permittee (i.e., collector/researcher). In order not to hinder research, curation, and education, the review should be completed as expeditiously as possible.

G. Dissemination of contextual data should be restricted only when there is a genuine risk to the collecting site. Restricting contextual data may affect the precision of research based on aggregated data, such as analysis of fossil occurrences in online public data portals. Therefore, restrictions should be imposed only if absolutely necessary, whereas all contextual data should be made available for research upon request.

H. Repository managers and other data providers should consider the needs of users for access to contextual data and other documentation when they evaluate sensitivity and weigh the impacts of disseminating data and restricting their access [note: for paleontological sites on U.S. Federal lands that fall under the U.S. Paleontological Resources Preservation Act, this determination is,

by law, the responsibility of the agency (permitter) that manages the land]. In cases where restrictions are placed on access to contextual data, the original data should be retained intact by the repository.

I. Original data should never be altered, falsified, or discarded.

J. Because research depends on the accuracy of data, users should be informed about omissions or changes that have been made to metadata in the interest of protecting a site. In cases where redacted data are disseminated, especially cases where the precision of geographic coordinates or stratigraphic placement has been purposefully reduced to protect the location of the collection site, the fact that this has been done should be distributed as part of the metadata for that specimen. In public databases, such as repository catalogs or data aggregators (e.g., online data portals), redacted records should be indicated with appropriate wording, rather than by leaving fields blank or null.

K. Users of contextual data should respect any and all restrictions that have been imposed by data providers. If granted enhanced access to restricted information, users must not compromise or otherwise infringe its confidentiality.

L. Whenever a data provider receives an application for access to restricted data, the assumption of continued sensitivity should be avoided. Rather, the occasion should be used as an opportunity to re-evaluate the determination. Decisions made by government agencies to release previously restricted contextual data must be made in consultation with the repository in order to meet the needs of non-governmental partners, the scientific community, and the general public. Cooperation with relevant governmental bodies is particularly important for repositories or situations where a 'freedom of information access' law applies in order to discuss potential ramifications of sharing requested sensitive information prior to its formal release.

III. Determining Sensitivity

Custodians of contextual data, including land managers, repository curators, repository collections staff, and data aggregators, are responsible for evaluating whether data in their possession should be regarded as ‘sensitive.’ Sensitive information is here defined as any contextual data, which, if released to the public, would create risk of harm to, or theft or destruction of, the resource or its origin (e.g., localities containing the paleontological resource), or adverse effect to a person or people. For vertebrate fossils, such potential harm most often stems from the possibility of damage to, or theft of, additional specimens of a rare taxon or additional parts of an individual organism that has previously been partially collected. Factors that should be taken into account when determining sensitivity include the type and level of threat, uniqueness of the taxon or attribute, type of information, and whether it is already publicly available. Although enumeration of all possible scenarios is impossible, it should be noted that non-paleontological factors may also make a site sensitive, including, but not limited to, environmental and ecological factors as well as anthropogenic (e.g., due to mining, development, or potential non-collecting visitors) and natural erosion rates. A criterion-based approach is outlined below as guidelines to assist in evaluating these factors.

A. Criteria for Site Vulnerability

An answer of 'yes' to any of the following criteria indicates potential of harm from human activity. The more 'yes' answers, the more likely the impact of harm there is.

1. Are there taxa that occur only at this site?
2. Do rare taxa occur at this site?
3. Does the site yield unusually complete or exceptionally well preserved fossils?
4. Is this site a unique or rare exemplar of its geological age or geographic provenance?
5. Is the site likely to produce additional specimens that will add significantly to the knowledge of variation within a taxon or to the completeness of a paleocommunity?
6. Are contextual field data that remain at the site easily damaged or destroyed?
7. Is the excavation or contextual documentation of the site still in progress?
8. Are scientifically important fossils that remain at the site easily damaged?
9. Is the site easily located because of proximity to roads or to easily visited or identifiable landmarks?
10. Are there non-paleontological resource values that would be impacted by releasing the location of the site?

B. Criteria for Determining Sensitivity

An answer of 'yes' to any of the following criteria indicates increased sensitivity of the data. The more 'yes' answers, the more likely the impact of harm there is.

1. Is the content and detail of the contextual data such that their release would enable someone to carry out activities that are harmful to fossils or data that remain at the site?
2. Would release of the contextual data bring them into the public domain for the first time?
3. Would release of the contextual data damage a partnership or relationship that is essential for preserving the site?
4. Would release of the contextual data increase traffic to the fossil site that would endanger non-paleontological resources, including endangered species, environmentally sensitive areas, or indigenous people cultural patrimony?
5. In cases where the contextual data do not include the precise location of a site, would its disclosure allow the location to be easily inferred from other publicly available information sources?

C. Recommendations for Georeferencing Based on Sensitivity Category

The above criteria should be used to assign a sensitivity category for sites using the following table. These criteria are intended to assist assessment, and are not prescriptive. Decision-making should incorporate the best available scientific principles and expertise. Geographic data should be disseminated with the following precision.

Category	Sensitivity	Georeference Guideline
1	Extreme	Location not released, or released only as broad categorical information (formation/region/county, etc.) without georeference coordinates
2	High	Georeference rounded to 0.1 degree (or nearest 1.5 km)
3	Medium	Georeference rounded to 0.01 degree (or nearest 0.5 km)
4	Low	Georeference rounded to 0.001 degree
5	Not sensitive	Unrestricted, full precision released

As a point of reference, 'medium sensitivity' (i.e., two decimal places in decimal degrees) roughly correspond to resolution in no greater detail than the nearest kilometer, or the last three places in Universal Transverse Mercator (UTM) coordinates. These guidelines are the same as used by the GBIF for sensitive localities

IV. Implementation for Repositories, Data Providers, Aggregators, and Distributors

Data custodians such as repositories or federal, state, and local governmental units should routinely consider providing complete unrestricted access to contextual data in their care unless these data have been assigned a sensitivity category of '2' or higher, as outlined above. When available, georeferenced coordinates for fossil occurrences should only be completely withheld for the most extremely sensitive sites. The following additional guidelines for dissemination of data should be followed.

A. Repository expectations

Data custodians are responsible for receiving, maintaining and preserving all contextual data related to localities, specimens, and collection acquisitions. While these data are maintained in public trust, complete access to data may be restricted at the discretion of the data custodian (for reasons outlined in Section III above). In the event that data are restricted, the repository manager should disclose this fact to data providers as well as data aggregators and distributors, or should include descriptive language to this effect on their respective online search forms. Should the extent of publicly available contextual data prove insufficient for a given purpose, data users are encouraged to contact individual repositories for more specific inquires. Repository managers should assess the needs of the user and the fitness for use of the request. Data users may be asked to provide the following justification to repository managers:

1. Name and institutional affiliation;
2. A description of the data they seek to obtain;
3. A description of their research, education, resource management, or other public benefit project, and why the requested data are pertinent or essential to their research questions; and
4. A description of how they intend to use and disseminate the data if the request is granted.

Repository managers are responsible for relaying institutional policies and specifying any terms and conditions that may be placed on information for release. It should be noted that contextual data are not necessarily always precise, accurate, complete, or reliable. Records may be unverified, vague, contain inherent errors, or reflect incorrect data. Data custodians should impress the importance of not using search results uncritically, as failing to acknowledge these limitations may undermine the legitimacy of certain data interpretations.

B. Data provider expectations

Data providers (e.g., authors) should bear ethical responsibility pertaining to any release of information stored in repositories and respect institutional policies and terms and conditions that may be placed on the information for release. If the information is permitted for release, data providers should provide georeferenced coordinates for fossil occurrences, preferably with maps and photographs, based on the sensitivity category of the site in question or the needs of the user

and the fitness for use. For fossils with ‘high’ to ‘extreme’ sensitivity, publicly accessible media (e.g., publications, public and conference presentations, social media, broadcast media, etc.) should not include images (photographs or video) containing landmarks that allow the sited to be easily located (e.g., highways, named/mapped geographic features, etc.). Authors should state whether the coordinates they provide are at full or reduced precision.

C. Data aggregator and distributer expectations

Data aggregators and distributers should continue to allow access to fossil site location data that are already in the public domain through scientific publications or occurrence databases and should add previously restricted occurrence coordinates whenever their sensitivity is deemed to have decreased to an appropriate level. Data aggregators must indicate via metadata those coordinates whose precision has been reduced when they redistribute georeference data. Data aggregators should note the date when the data were last updated and the source from where they were obtained.