
CONSOLIDATED METADATA FOR THE dbSEABED DATA HOLDINGS

An Index of all the Contributing Data Sources

Version: 19 Dec 2019

THE METADATA LISTING:

To see the current whole metadata listing open
"[http://instaar.colorado.edu/~jenkinsc/dbseabed/source_metadata/dbSBD
AllSourcesOfData.txt](http://instaar.colorado.edu/~jenkinsc/dbseabed/source_metadata/dbSBDAllSourcesOfData.txt)" (20MB)
while referring to the notes below.

NOTES:

A. The data volume this metadata describes holds >9,000 datasets, >4 million sites, and >6 million samples. Such a large amount of parameter data is required to adequately describe the worldwide diversity of seafloor materials.

B. The metadata set will change over time as datasets and metadata are added, improved, or even decommissioned (if found faulty). Note: relational database keys are assigned to only when data is integrated to database- and GIS-ready format.

C. The metadata consist of (i) 'active' metadata (fields 1-17), (ii) 'inactive' metadata (fields 18-20). The former are controlled in format and vocabulary, and are suitable for databasing. The second are uncontrolled, information-rich texts for reading and mining.

D. The metadata are not complete in every respect. This is because dbSEABED does not put in (speculate on) an entry where the original authors provided none. A gap remains which is marked by a blank field, '##', '--', 'NA', or '?' in the metadata.

E. Further embedded metadata exists for the datasubsets, sites, samples, horizons, subsamples, but is not given here.

F. Despite best efforts by all the contributors and aggregators, the metadata are not guaranteed to be free of errors. However, the error rate is exceedingly low. For continual improvement users should notify dbSEABED of any issues or errors that are detected in the 'active' metadata. Some errors in the 'inactive' data fields are likely to have

been carried directly from the original sources, or are from the transcription process (e.g., from PDF encodings).

THE FIELDS:

activeFlag Flag for how actionable a data line is (e.g.

'SRC' Marker in the data pool for source information

DataSetMnemonic Brief name identifier for each dataset, a mnemonic.

(Usually a concatenation of author, year, several keywords.

DataOwner The primary organization, entity which collected, composed, served, or owns the data.

DataPerson A primary person responsible for collecting, assembling, describing, or providing the dataset.

SourceProtectionLevel Level of restriction, if any, on the whole original document or dataset (e.g. 'copyright', 'confidential', 'public', 'unencumbered').

ReleaseProtectionLevel Level of restriction, of any, on the selected, edited, reformatted subset of data which is held and served from dbSEABED (e.g., 'restricted', 'public', 'unencumbered').

AttributeDataEntryMethod Whether the attribute data (observations) were entered from free text, a table, graphic, map, database, or web-served data (e.g., 'editText', 'editTable', 'digiMap')

PositionEntryMethod Whether the locating data (positioning) were entered from free text, a table, graphic, map, database, or web-served data (e.g., 'editText', 'editTable', 'digiMap')

DataSourceType The type of source publication, document, dataset (e.g., publPaper, publReport, unpublMScThesis, database, gisShapefile, webData).

Region Brief description of the source geographic region of the data (e.g., "North Sea").

SurveyDate Date of the survey, collection, observation of the data - where known. This is loose format, but may be accompanied by a pattern (e.g., 'ddMMMy Z') to assist decoding.

ReportDate Date of the earliest record of the data (e.g., date first submitted for publication, date given in the document properties). This is loose format, but may be accompanied by a pattern (e.g., 'ddMMMy Z') to assist decoding.

NavMethod An indication of the technology used for positioning (e.g., 'LORAN', 'DGPS', 'SatNav', 'Radar fixes', 'Dead reckoning', 'Celestial').

EntryDate Date of entry into dbSEABED.

DataTypes The attributes which are given in the data (e.g., porosity, carbonate, grainsizes, color).

DOI A Digital Object Identifier, if available, for the publication or the data itself.

datasetSrcText Concatenated metadata describing the source of the dataset, authors, institution, publication, dates, filename, supplementary data filename, etc. Will contain a formal citation where available. (Lines delimited with '***'.)

sitesSrcText Concatenated metadata describing the location, site selection, positioning methods, sampling / observing methods and gear, subbottom penetration, subsampling, etc. (Lines delimited with '***'.)

otherSrcText Concatenated metadata describing all other aspects of the data, the described and measured attributes, methods of analysis and reporting, calculation of other variables. (Lines delimited with '***'.)