

Guidelines for Metadata Review of Data

Metadata records are generally best reviewed alongside the data, since metadata describe datasets and there are some points at which it will be helpful to compare what is stated in the metadata to what the dataset shows. A data product may require more than one metadata record – for example, a record describing a collection of data sets, or records for individual datasets.

The role of the metadata reviewer is to ensure the accuracy, completeness, and usability of the metadata for data products to be approved and released by the USGS.

If the metadata reviewer is also serving as data reviewer, see the additional guidelines for data review (<http://www.usgs.gov/datamanagement/share/datarelease.php>).

The following checklist is provided for the assistance of metadata reviewers who are familiar with metadata standards, rather than as training for beginners.

Review Process Checklist:

Check compliance with standards using a recommended metadata validation tool (<http://www.usgs.gov/datamanagement/describe/metadata.php>). Currently, USGS recommends using the Metadata Parser to validate the FGDC Content Standard for Digital Geospatial Metadata standard.

Make note of any compliance issues from the error report issued by the validation tool. (Examples of errors include dates in non-compliant format, omission of required information, etc.)

Because validation tools are unable to check the quality of information in the metadata, perform quality checks on the metadata to confirm that:

- the metadata matches the data and its correct version (e.g., is this metadata record describing the correct dataset? Sometimes existing metadata records are used as templates and some old information carries over to the new record);
- links to data, publications, and services exist (although links may change when final DOI is assigned);
- metadata explains field names and values contained in the dataset;
- if applicable, geographic coordinates match location keywords in metadata and these agree with the data;
- keywords accurately represent the data and include terms from standard vocabularies whenever possible (such as the USGS Thesaurus, <http://www.usgs.gov/science/services.html> or Biocomplexity Thesaurus, http://www.usgs.gov/core_science_systems/csas/biocomplexity_thesaurus/);
- information about data processing steps, methodology, and lineage are included in the record and match any associated publications;

- the metadata provides complete and current information about how to use the data files – access instructions, unusual software requirements, data models, definitions of terms, size of the data, etc;
- access constraints, use constraints and distribution liability statements are included and reflect USGS policy;
- content is written in a way that is usable and helpful, without use of formatting or special characters that will fail to be transferred to XML. (See metadata best practices on the data management website, www.usgs.gov/datamanagement/describe/metadata.php)

Once the review is complete, return comments and suggestions to the metadata author for updates and improvements.