Guidelines for Metadata Review

A data product may require more than one metadata record – for example, a metadata record for each dataset in a collection of datasets, and another metadata record for the collection as a whole. Metadata records are best reviewed alongside the data, because metadata describe data and there are some points at which it will be helpful to compare what is stated in the metadata record to what the dataset actually includes.

It is helpful if associated science products (reports or publications) are available for reference during metadata review.

The role of the metadata reviewer is to ensure the accuracy, completeness, and usefulness of the metadata for USGS data products. If the metadata reviewer is also serving as data reviewer, see the additional guidelines for data review (https://www2.usgs.gov/datamanagement/share/datarelease.php).

The following checklist is provided for the assistance of metadata reviewers who are familiar with metadata standards. The checklist provides general guidelines and can be customized, as necessary, for use in specific USGS programs or science centers.

Review Process Checklist:

Check compliance with the Federal Geographic Data Committee (FGDC) Content Standard for Digital Geospatial Metadata (CSDGM) or International Organization for Standardization (ISO) standards using a recommended metadata validation tool (https://www2.usgs.gov/datamanagement/describe/metadata.php). Currently, USGS recommends using the Metadata Parser (https://geo-nsdi.er.usgs.gov/validation/) to validate the FGDC CSDGM 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 the information. For example, verify that:
- the title conveys the "what, where, when, who, and (if applicable) scale" of the data;
- the abstract provides a short statement that allows users to quickly evaluate the content of the dataset;
- if applicable, the coordinate system and datum are defined appropriately (both horizontal and vertical);

- 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 function properly (links that are not yet known or defined should have clearly marked placeholders that will be replaced later);
- data tables, fields, and values must be explained clearly so that a non-specialist can understand them; verify that entity and attribute content of metadata corresponds to data set;
- as warranted, geographic coordinates are provided, and they match location keywords in metadata and also agree with the data;
- keywords accurately represent the data and include terms from standard vocabularies whenever possible (such as the USGS Thesaurus, https://www2.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; in particular:
  - could a scientist or technician recreate the final data set from the descriptions?
  - can documentation about methodology be easily found and used?
  - are processing software and versions identified?
- 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;
- the data quality section adequately describes the procedures used to ensure the quality of the data;
- access constraints, use constraints, and distribution liability statements are included and are consistent with USGS policy (see: https://www2.usgs.gov/fsp/fsp_disclaimers.asp);
- if the data or metadata has been revised, correct version identification is provided (see: https://www2.usgs.gov/fsp/data_versioning.asp);
- content is written in a way that is usable and helpful, without use of special formatting or characters that will fail to be transferred to XML (generally metadata should be encoded using UTF-8).

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