Twenty Questions about Geospatial Data
(to ask in preparation for writing formal metadata)

https://njgin.state.nj.us/oit/gis/NJ_NJGINExplorer/docs/MD_TwentyQuestions.doc
https://njgin.state.nj.us/oit/gis/NJ_NJGINExplorer/docs/MD_TwentyQuestions.pdf

1) What is the subject of the data set? Be as specific and succinct as possible. Is there a file name by which the data set generally is known and/or distributed?

2) What is the approximate geographic location of the project? Please provide a general description such as a state plus county and/or city name, watershed management area name, or other general locator.

3) Who is the originator/owner/creator of the data set?
   a. Who (organization and/or person) defined the data project?

   b. Who actually did the work to create the data set?

   c. Who should be listed to answer questions about the data set, if not the metadata author? What is their full contact information?

   d. Are there other organizations and/or individuals who should get credit for support, funding, data creation? Was a vendor used to collect and/or analyze the data?

OGIS
NEW JERSEY

4) Briefly describe the data set, in a few sentences that are understandable to an intelligent non-specialist. Include spatial data type (point, line, polygon, raster, tabular coordinates only, etc) and format.

5) Does the data set have a Spatial Reference assigned to it? If not, what is the projection and datum, or coordinate system, of the data set? If a spatial reference is not already attached to the data set, try to provide a correct EPSG reference from the Web site published by the International Association of Oil and Gas Producers (http://www.epsg-registry.org/)

6) Why were the data collected or created – i.e. what was the business purpose? How will the data be used or how can they be used by others?

7) What time period is represented by the data set? Is it known when the data matched the real world (e.g. photo capture dates)? Was the data set drawn from published sources with publication dates, or downloaded from a Web site on a date or dates?

8) What organization has/will publish the data set? Where is that organization located? When was or will the data set be published, or re-published?

9) In what state is the data set – complete, in progress, or planned? If complete or in progress, are there plans to update the data set? If so, how often?

10) What are some subject and geography keywords or phrases that might be used to search for the data set? If the subject area of the data set has a standard language list (Thesaurus), list it and the appropriate keyword(s). If the subject matter is Geology, Soils, Archaeology, History, or the like, consider specialized keywords.

11) What processing was done to create or modify the data? Were specific models or tools used? Can the processing be described in a sentence or two, or is there more necessary detail? On what date(s) was the processing performed or finished?

12) Were source data sets or publications used? If so, for each, provide (where available):
    a. Source Title, Originator/Author/Editor, Publication Date
    b. Source Time Period of Content
    c. Source scale (paper map) or nominal scale if known (digital data)
    d. Source format and media, file name
    e. How the source contributed to creating the data set being documented

OGIS
NEW JERSEY

13) Was the positional accuracy of the data assessed? Was any testing done? If not, can an estimate be described based on the sources? Was any topology testing/repair done?

14) Are there any there gaps, exclusions, missed areas, or changes in data collection in time or space?

15) If a Data Dictionary has been created, provide it and expand, if necessary, to include the below requested information. Otherwise describe in narrative:

   a. Provide a description of each field in the attribute table, except for those automatically generated by software. The description should provide a label, an alias if different, a definition of the field contents, and list the organization and/or person who defined it.

b. Information also is needed about the permissible contents of each field. If the contents of a field are confined to a published set of codes, provide reference information for the code list (preferably with URL.) If the contents are restricted to a list defined only in this data set, provide the list and the meaning of each entry. If there is a theoretical range in which numerical values should fall, provide that. If the contents cannot easily be described in any of these ways (e.g. a list of people's names), just describe the genesis of the entries.

16) Was any quality checking of the data set attributes done? If so, describe.

17) If the data are being distributed:
    a. Who is the primary contact for the most important distribution method? What is the complete contact information?

    b. In what format are the data being distributed (e.g. shapefile, file geodatabase, kmz, WMS, WFS, etc.)

    c. Is there any cost to acquire or access the data?

    d. What is an actionable URL for (in order of desirability): live online service; directly-downloadable package; searchable catalog; information Web page. List any that are available.

18) Is distribution or use of the data or metadata restricted in some way? Access? Use? Security Clearance? Licensing? Are the data Survey data (created by and certified by a licensed land surveyor?) Is there a legal liability disclaimer?

19) Is the data set in support of a larger published work, or part of an ongoing formal series? If so, provide Citation (bibliographic) Information for that work or series – Title, Originator/Author, Publication Place, Publication Date, Publisher, Series/Issue identification information, etc.

20) Is there anything else that should be recorded about the data, or advice for potential users? This might be additional technical details, and/or critical information that is elsewhere in the metadata but might be missed.

❖❖❖

The list above assumes that the metadata writer will have access to a copy of the data during creation of the formal metadata, allowing automatic capture of some information from the data.

Questions about this document may be directed to:
Edith Konopka, NJ Office of Information Technology, Office of Geographic Information Systems, edith.konopka@tech.nj.gov
OR
The New Jersey Geographic Information Network administrator: njgin@oit.nj.gov

Content of this document was drawn in part from previous work by:

Nina Savar, Northeast Illinois Planning Commission;
John Bocchino, New Jersey Department of Environmental Protection;
Peter Schweitzer, U.S. Geological Survey;
U.S. Geological Survey – Data Management Guidance Materials/White Papers, USGS Data Management website http://www.usgs.gov/datamanagement/index.php, November 2015.