

[Comments by Greg Smith, California Department of Water Resources.](#)

[March 16, 2012](#)

[The idea of minimum metadata standards is a good idea. However, the minimum standards lack any context or discussion. There is a lot more than just this list that needs to be covered. Absent topics include:](#)

1. [To what class of spatial data would these standards apply? Legacy data? To historical data sets that are still used? To any data set developed after a certain date?](#)
2. [To whom do these standards apply? To just State agencies? To data from outside of the State? To data created by local and regional governments?](#)
3. [Would these standards apply to the non-spatial data sets, say something in Oracle? Would these standards apply to the spatial elements of data sets? What happens to other fields in a data set, say for instance a license issued by the State of California? Would this field have to be described the way that spatial data fields would have to be described?](#)
4. [What happens if a data set does not meet these minimum standards? Historical, legacy and data from outside of California may not. What then? We do not publicize the data? We do not include it in the State's GeoPortal?](#)
5. [Who will be responsible for completing the metadata? Again, this is a question for historical, legacy and data from outside of California.](#)
6. [Do the standards apply equally to government employees and contractors?](#)

[Another necessary piece of information is how each of the items listed below matches up with the FGDC guidelines. When you identify the contact information, do you want all the mandatory contact information identified by the FGDC? Do you want all the information \(mandatory, conditional and applicable\)? Of do you just want the information identified here, regardless of what the FGDC standards say? \(For instance, the minimum standards are missing address, city, state and zip code. These would be mandatory under the FGDC standards.\)](#)

[Because we are trying to define a minimum, and not a robust, metadata standard. I would consider taking out field definitions and abbreviation definitions. Yes, these are great to know. But if the information is not there, do we reject the metadata? I would prefer to have the metadata without this information and know the data set exists.](#)

-

MINIMUMS DETAIL

Abstract: Briefly describe what the data set is about (who, what, where, when). Include any limitations of the dataset, assumptions made, and if there is anything special that the user of these data should be aware of.

Name: uniquely identifies the dataset.

Purpose: Briefly describe why the data set was created.

Date: The date or range of dates when the data were gathered, or the date the photos, maps or other items at the core of the data set, were created.

Point of Contact: Contact information for an individual or organization that is knowledgeable about the data set. Include:

Person's Name: Complete first and last name

Organization's Name: Program, administrative unit, and agency, company, or group name

Telephone Number: Including Area Code

E-Mail address:

Field Definitions: ~~List and d~~Define each field, including the data type. If you have multiple tables, then include the table name for each field. Where field names may be similar to field names generated by GIS software, such as length or area, identify if this is a field defined by the user or the software.

Abbreviation Definitions: For any field that contains numeric or alphabetic codes (e.g., SAC = Sacramento County), list each code/abbreviation and provide an unabbreviated definition.

Access Constraints: Is there a need to limit who has access to see or read this dataset? If so, specify. If not, put "None".

~~**Use Constraints:** Is there a need to limit the use of this dataset to certain people or to specific tasks? If so, specify. If not, put "None".~~

~~**Citation:** (optional) Also include how the data should be cited, if you want something specific.~~

~~**Distribution:** Define distribution constraints. Also designate the location of the data. If the data is distributed as a web service or end point, provide the url or link to the service~~

~~**Progress:** Complete or Incomplete. What about something that may be continually in progress? The location of new wells or roads, for instance.~~

Update Frequency: Possible values are: Continually, Daily, Weekly, Monthly, Annually, Unknown, As Needed, Irregular, None Planned, or

Projection: What is the Projected Coordinate System name?

Please define the complete projective information for your data here. I think the Californian Technology Agency should define a few of the more popular projections. A person could the reference one of those definitions in the metadata, without having to specify a complete projection, according to the FGDC.

Datum (or Geographic Coordinate System): Which Datum is the projection in?

NAD83 (GCS_North_American_1983) (preferred)

NAD27 (GCS_North_American_1927)

WGS84 (WGS_1984)

ISO Topic Category an Category Code (one or more of the following):

Farming	001
Biota	002
Boundaries	003
Climatology/ Meteorology/ Atmosphere	004
Economy	005
Elevation	006
Environment	007
Geoscientific Information	008
Health	009
Imagery/BaseMaps/ EarthCover	010
Intelligence/Military	011
Inland Waters	012
Location	013
Oceans	014

Planning / Cadastre	015
Society	016
Structure	017
Transportation	018
Utilities / Communication	019

Keywords: keyword tags that define the data. Example: Forest Cover – trees, canopy, woodland, coniferous, etc...

I think the Californian Technology Agency should develop a standard list of keywords for spatial data. At a minimum, the metadata would include any of those relevant keywords. The metadata could contain more keywords defined by the data steward or the steward's organization. The keywords for CERES would be a good place to start.