

# FGDC metadata and multiple versions of datasets

I'm looking for feedback. The NOAA climate data portal uses FGDC metadata as the only method to catalog files. We archive datasets at ESRL that are created elsewhere (and in a few cases, someone else archives a dataset created here).

NOAA would like one version of FGDC metadata for each dataset. This is fine if one discusses how background of the dataset, how it was put together, areal and temporal extents and similar info. However, FGDC metadata also has info on how to read and download files. This would obviously be an issue if the data were at 2 or more places, possible in different formats as some datasets would not easily be found and the use data might be confusing. But, having multiple FGDC metadata would mean it would be hard for us to include all the correct background info (and needless to say, we would not know when it was changed at the source). How can we reconcile the need to have one version of dataset background information yet multiple versions of how to read and get the data?

Complicating this is that we could have products created from data we archive (say, climatologies). Should we have separate FGDC metadata?

That doesn't totally make sense yet if we don't, our products are not findable. Or, we could have a subset of the whole dataset. If we point back to the originator's FGDC it might be confusing as we don't have all the variables.

Note- dataset for me refers to a collection. Say, all the NCEP/NCAR R1 products or all the CMAP precipitation products. NOT single files or variables.

We will be discussing with NCDC soon and I'd like to have a solution that leaves our products in their catalog so any feedback is helpful.

Cathy