

## Member Nodes - Task #7461

MNDeployment # 7082 (Operational): USGS Science Data Catalog (SDC)

### Determine how to handle metadata which is well-formed but does not validate

2015-10-30 16:05 - Bruce Wilson

<b>Status:</b> Closed	<b>Start date:</b> 2015-10-30
<b>Priority:</b> Normal	<b>Due date:</b>
<b>Assignee:</b> Dave Vieglais	<b>% Done:</b> 100%
<b>Category:</b>	<b>Estimated time:</b> 0.00 hour
<b>Target version:</b> Operational	
<b>Story Points:</b>	
<b>Description</b> <p>SDC is an example of cases where there are metadata records that are declared as FGDC, but contain extensions (such as the ESRI extensions) so that the file does not actually validate according to the declared schema. The objective of this task is to determine which approach to take. One extreme is the current approach, where only metadata files that pass validation are accepted. Another is to allow any well-formed XML (and potentially set the sysmeta to declare the metadata type as text/xml). From Bruce's perspective, if we accept metadata that's not schema valid, it should be straightforward to determine from sysmeta (or something similar) which metadata is schema valid and which is not. There are intermediate proposals along the idea of looking to see if schema invalid metadata matches one of a list of known alternatives and only accepting the metadata where it matches something from a known list of variants to the declared schema.</p> <p>The output from this task is a plan with tasks that will at least allow the ESRI-extended FGDC metadata from SDC to be accepted into DataONE.</p> <p>Likely requirements:</p> <ul style="list-style-type: none"><li>* The metadata must at least be well-formed XML</li><li>* The metadata must be at least interpretable for discovery.</li><li>* Tools need to be able to determine if the metadata is valid for those tools that want to do deeper interpretation (like the Matlab or R tools)</li><li>* It is useful to be able to provide feedback to MNs for the metadata that's interpretable but not schema valid.</li></ul>	
<b>Related issues:</b>	
Related to Member Nodes - MNDeployment #3238: Idaho Northwest Knowledge Netwo...	<b>Operational</b> <b>2013-01-30</b>
Blocks Member Nodes - Task #7083: USGS SDC metadata issues with CSDGM	<b>Closed</b> <b>2015-05-07</b>

### History

#### #1 - 2015-11-02 19:13 - Bruce Wilson

- Blocks Task #7083: USGS SDC metadata issues with CSDGM added

#### #2 - 2016-06-17 19:40 - Laura Moyers

- Related to MNDeployment #3238: Idaho Northwest Knowledge Network member node added

#### #3 - 2018-04-27 20:29 - Amy Forrester

- % Done changed from 30 to 100

- Status changed from In Progress to Closed

per Dave - 4/27/2018