Member Nodes - Story #8819

IEDA documents not in DataONE

2019-06-13 19:20 - John Evans

Status:	New		Start date:	2019-06-13
Priority:	Normal		Due date:	
Assignee:			% Done:	0%
Category:			Estimated time:	0.00 hour
Target version:				
Story Points:				

Description

While testing schema_org adapters for ARM and IEDA, I noticed that 9 documents found in the IEDA site map have malformed JSON-LD script elements. These nine documents do not appear in DataOne when using search. There are two issues arising in the malformed JSON-LD:

1. description values have over-escaped double-quotes, i.e. \"bipolar seesaw\" rather than \"bipolar seesaw\". See IEDA documents with IDs 601015, 601098, e.g. http://get.iedadata.org/metadata/iso/601015
2. description values have embedded double-quotes that are not properly escaped at all. See IEDA documents with IDs 600165, 601033, 601076, 601077, 601089, 601103, 601178

Fixing the over-escaped double-quotes is easy to do on the fly, but the under-escaped double-quotes requires taking a bit more care (still doable on the fly).